AD A042075

ARI TECHNICAL REPORT
TR-77-A11

DEVELOPMENT AND EVALUATION OF A CONTENT ANALYTIC APPROACH IN ARMY FIELD SYSTEM DATA ORGANIZATION

Arthur I. Siegel, Mark G. Pfeiffer, Tom M. Warms

APPLIED PSYCHOLOGICAL SERVICES, INC.
Science Center
Wayne, Pennsylvania 19087

AUGUST 1977

DDC
RECEIVED
JUL 26 1977
D

Contract DAHC 19-75-C-0037

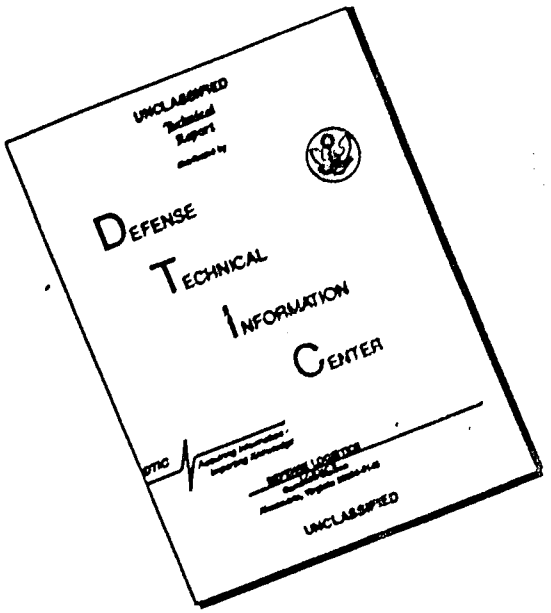
Prepared for

U.S. ARMY RESEARCH INSTITUTE
for the BEHAVIORAL and SOCIAL SCIENCES
5001 Eisenhower Avenue
Alexandria, Virginia 22333

AU No.
DDC FILE COPY

Approved for public release; distribution unlimited.

DISCLAIMER NOTICE



THIS DOCUMENT IS BEST QUALITY AVAILABLE. THE COPY FURNISHED TO DTIC CONTAINED A SIGNIFICANT NUMBER OF PAGES WHICH DO NOT REPRODUCE LEGIBLY.

U. S. ARMY RESEARCH INSTITUTE
FOR THE BEHAVIORAL AND SOCIAL SCIENCES

A Field Operating Agency under the Jurisdiction of the Deputy Chief of Staff for Personnel

J. E. UHLANER
Technical Director

W. C. MAUS
COL, GS
Commander

ACCESSION for
NTIS White Section
DDC Buff Section
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
Dist. AVAIL. and/or SPECIAL
A

D D C
RECEIVED
JUL 26 1977
D

Research accomplished
under contract to the Department of the Army

Applied Psychological Services, Inc.

NOTICES

DISTRIBUTION: Primary distribution of this report has been made by ARI. Please address correspondence concerning distribution of reports to: U. S. Army Research Institute for the Behavioral and Social Sciences, ATTN: PERI-P, 5001 Eisenhower Avenue, Alexandria, Virginia 22333.

FINAL DISPOSITION: This report may be destroyed when it is no longer needed. Please do not return it to the U. S. Army Research Institute for the Behavioral and Social Sciences.

NOTE: The findings in this report are not to be construed as an official Department of the Army position, unless so designated by other authorized documents.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(18) ARI (19) TR-77-A11

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report TR-77-A11	2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER	
4. TITLE (and Subtitle) (6): DEVELOPMENT AND EVALUATION OF A CONTENT ANALYTIC APPROACH TO ARMY FIELD SYSTEM ORGANIZATION	5. TYPE OF REPORT & PERIOD COVERED (9): Final rept.,
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10): Arthur I. Siegel, Mark G. Pfeiffer and Tom M. Warms	8. CONTRACT OR GRANT NUMBER(s) (15): DAHC 19-75-C-0037
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Applied Psychological Services, Inc. Science Center Wayne, Pennsylvania 19087	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS (16): 2Q762722A765
11. CONTROLLING OFFICE NAME AND ADDRESS: US Army Research Institute for the Behavioral and Social Sciences, 5001 Eisenhower Avenue, Alexandria, Virginia 22333 PERI-OS	12. REPORT DATE (11): August 1977
	13. NUMBER OF PAGES: 209 (12) 219 p.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report): Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Research monitored technically by Dr. Stanley M. Halpin of the Battlefield Information Systems Technical Area of the Army Research Institute.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Information systems	Multidimensional scaling	Content analysis
Linguistics	Field language	Taxonomies
Psycholinguistics	Battlefield intelligence	
Message classification	Information organization	

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The logic, methods, and results of a study into the derivation of a content analytic approach to Army field system data organization are described. The first steps of the program involved a linguistic analysis of a set of battlefield messages and multidimensional scaling analysis of a sample of messages representing all information classes yielded by the linguistic analysis. Four multidimensional scaling analyses were completed.

402 774

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

20.

Each was based on the linguistic similarity perceptions of one of four experienced battlefield analysts. High agreement was found among the factorial structures yielded by the data of each battlefield analyst. Accordingly, an overall analysis was completed, and 15 factors were identified as representing the perceptual substrate of the Army field information linguistic system. On the basis of the derived factors, a battlefield language taxonomy was developed. The taxonomy was tested in two separate field oriented experiments. The results of these experiments indicated that intelligence analysts can classify messages reliably within the taxonomy and that they can reliably use the taxonomy for inquiry purposes. Moreover, adequate subjective confidence was expressed by the analysts in the use of the system. Finally, a computer system for automatic classification of battlefield messages is presented.

ABSTRACT

The logic, methods, and results of a study into the derivation of a content analytic approach to Army field system data organization are described. The first steps of the program involved a linguistic analysis of a set of battlefield messages and multidimensional scaling analysis of a sample of messages representing all information classes yielded by the linguistic analysis. Four multidimensional scaling analyses were completed. Each was based on the linguistic similarity perceptions of one of four experienced battlefield analysts. High agreement was found among the factorial structures yielded by the data of each battlefield analyst. Accordingly, an overall analysis was completed, and 15 factors were identified as representing the perceptual substrate of the Army field information linguistic system. On the basis of the derived factors, a battlefield language taxonomy was developed. The taxonomy was tested in two separate field oriented experiments. The results of these experiments indicated that intelligence analysts can classify messages reliably within the taxonomy and that they can reliably use the taxonomy for inquiry purposes. Moreover, adequate subjective confidence was expressed by the analysts in the use of the system. Finally, a computer system for automatic classification of battlefield messages is presented.

TABLE OF CONTENTS

PRECEDING PAGE BLANK-NOT FILMED

TABLE OF CONTENTS (cont.)

TABLE OF CONTENTS (cont.)

LIST OF FIGURES

LIST OF TABLES

LIST OF TABLES (cont.)

LIST OF TABLES (cont.)

LIST OF TABLES (cont.)

LIST OF TABLES (cont.)

CHAPTER I

INTRODUCTION

Computer based Army tactical intelligence systems must consider a wide variety of message types involving a multiplicity of content areas. Moreover, such diversified information must be entered accurately into a data information bank and be in such a form that the information can be retrieved quickly, thoroughly, and accurately. Messages received for entry into the system may range from informal reports relative to the morale and motivation of hostile forces to "hard" information about troop movements and armament. In order that such information will be entered in a systematic manner, a standard scheme is required. Such a scheme would allow not only quick and accurate data entry but also data recovery. The present study had the goal of developing a system for the classification and retrieval of battlefield messages based on linguistic principles and language perceptions of the user.

Objective

The specific purposes of the present program were to develop an approach to the organization of the linguistic data which are employed in an Army battlefield oriented, computer based intelligence system. To achieve this goal, a sequential set of steps was involved: (1) content and factor analysis (multidimensional scaling) to develop a basis for organizing and classifying such linguistic data, (2) development of a taxonomy for categorizing such data, (3) test of the utility and reliability of the organizational/classificatory taxonomy, and (4) preliminary definition of a system for automatic classification of battlefield messages.

Linguistic Logic of Approach

The basic linguistic logic of the present approach was that the developed classification method, if it is to be at all useful, must be based on a substrate which mimics intelligence analyst language processing and understanding in the tactical intelligence situation. We hold that schemes which are not compatible with these processes will fail to possess utility regardless of their merits in other regards.

Additionally, if the same information appears in two different forms, as in: "The enemy was seen building up troops," and "An enemy troup buildup was seen," the scheme must recognize the redundancy. More importantly, in a message such as "I saw enemy troup movements, and a captured enemy soldier reports that there are additional troops at a distance of 10 kilometers," the method must permit a differentiation between what the sender of the message was able to report as being true, and what was found out from a possibly unreliable source. The message might then be classified either on the basis of the reliable or on the basis of the unreliable information, or on both.

A method which depends on picking out key words or phrases as a basis for classification, on the other hand, cannot make such a subtle distinction. Since the message contains references to a captured soldier, any classificatory scheme which uses a "captured soldier" key word will result in the entire message being associated with that keyword. Similarly, if the classificatory scheme does not use a "captured soldier" keyword, then the fact is lost that a captured soldier is a source for some of the information in the message.

The requirement for such subtle distinctions is even more important in the present context because there is a greater permanence to the classification scheme than heretofore. A description of the elements which are emerging in a particular situation is perhaps enough to classify messages when the user is fully aware of the total situational context. However, in the present case, this intimate familiarity cannot be assumed, and the relationships among emerging elements becomes crucial.

Transformational Grammar

The methodology of transformational grammar developed by Harris (1968) provided the basis for a grammar of the battlefield situation. The nature of the grammar arrived at, when one is dealing with a natural language like English, is discussed in detail in Harris' Mathematical Structures of Language. Similar work has been done on other natural languages, most notably, the work of Gross (1968) on French. The work of Hiz and Joshi (1968) is also pertinent. In the present work, these same techniques were used to deal with the battlefield intelligence situation. The language of the battlefield situation is of course English, but it differs from ordinary English in that grammatical forms in the battlefield situation would not be considered grammatical in ordinary English. For example, in the battlefield situation, the report "Enemy buildup between 0800 and 0900" is acceptable. In ordinary English, this is not an acceptable sentence.

In this respect, battlefield English is similar to "telegraph English." In addition, battlefield English is much smaller in extent than ordinary English. There are many words and many methods of combining words into sentences which are acceptable in ordinary English that would never be encountered in battlefield English. For example, one would not expect to find the word "tumultuous" at the battlefield, nor the form in which the object of the sentence is brought to the beginning of the sentence, as in "This I know." Accordingly, in a sense, battlefield English can be considered a sublanguage of English. Harris dealt with sublanguages as follows:

> Thus the sublanguage grammar contains rules which the language violates and the language grammar contains rules which the sublanguage never meets. It follows that while the sentences of such science object-languages are included in the language as a whole, the grammar of these sublanguages intersects (rather than is included in) the grammar of the language as a whole (p.155).

According to the Harris theory, every language is characterized by classes of operators and kernels, such that each sentence can be written as a product of some number of them. The operator and kernel classes differ from language to language. Examples of kernel forms for English are:

noun, verb, noun	The boy hits the girl.
noun, verb, noun, preposition, noun	The boy throws the ball to the men.

The operator types are:

Word expansion
    The flower is very pretty.
Verb operators
    I enjoy going to parties.
Sentence operators
    I suggest you go home.
Connectives
    The men who objected called the police.
Permutations
    This I like.
Zeroings and pro-wordings
    It is an interesting book.
Morphophenomics (which give the appropriate final shapes to the resultants of the other operators)

One has various options available for dealing with a sublanguage of English. One of the possibilities is to start with a grammar of English, and make additions and deletions where necessary. A second is to apply a grammar of English directly to the corpus under consideration. A third is to use the same principles that are used in developing a grammar of English, but only to consider the facts of the sublanguage itself, without reference to the facts of the entire English language. It was believed that the third approach would arrive at a defensible set of results in the most efficient manner. Accordingly, the battlefield sublanguage, without reference to the total language, formed the basis for the present work

Psychometric Considerations

While a formal linguistic analytic system will tell us something about the global foundation of the language involved, it will tell us little about the perceptual organization of the language in the mind of the language user. A grammar is extrinsic to the human; the structure of the language in the mind of the user is a separate issue. In fact, the perceptual structure may vary among user subpopulations. To derive the perceptual structure of battlefield analysts relative to battlefield English, multidimensional scaling techniques were employed in the present work.

Multidimensional scaling and related clustering techniques are concerned primarily with the spatial representation of relationships among behavioral data--in our case, similarity perceptions by experts among military field information statements.

In areas where the variables are complex and the dimensions unknown or doubtful, it is appropriate to delineate the variables through multidimensional scaling analysis, rather than to establish the dimensions arbitrarily. Evidence of the validity of the multidimensional method has been cited, especially in research where dimensions are well established.

The utility of the multidimensional scaling for yielding an empirically sound and logical basis for describing a complex has been demonstrated in a variety of contexts including attitudes (Messick, 1954, 1956; Abelson, 1954), personality (Jackson, Messick, & Solley, 1957), jobs (Reeb, 1959), civil defense (Smith & Siegel, 1967), and display evaluation (Silver, Landis, & Jones, 1965), Naval electronic maintenance (Schultz & Siegel, 1964), Air Force tasks (Siegel & Bergman, 1973), and marketing analysis (Green & Carmone, 1970). Messick, in his studies on color, concluded that "since multidimensional scaling procedures yielded structures which correlated highly with the revised Munsell system, it would now seem reasonable to apply these procedures for purposes of exploration and discovery in areas of unknown dimensionality" [1956, p. 374].

While it is recognized that the methodological issues of multidimensional scaling continue to remain somewhat fluid (Coombs et al., 1970), the importance of these theoretical controversies is not considered to be highly relevant to the present study. Actually, any of the content or distance methods would probably result in a reasonable and usable structure (Sjoeberg, 1975).

In the content approach (Ekman, 1965), judges provide estimates of the similarity between all possible pairs of the stimuli under consideration. The similarity magnitude estimates are made along a 0 to 100 scale, and these values may be taken as correlation coefficients. A large amount of work has been done with the Ekman similarity analysis technique and how it compares with other multidimensional scaling techniques since Ekman first introduced the technique in 1954 (Andrews & Ray, 1957; Brown, 1967; Coombs, 1964; Ekman, 1954, 1965, 1969; Kuennapas, 1966, 1967; Pfeiffer, Kuennapas, & Fastiggi, 1973; Sjoeberg, 1975; Stone & Coles, 1971). Ekman (1970) concluded that such content methods are superior to distance methods when the purpose is to reveal underlying dimensionality in such areas as color perception, odors, emotions, and personality traits. While Kuennapas (1966, 1967) found that some mechanism other than similarity was at work in the perception of form and memory of form, the similarity analysis revealed the underlying dimensionality. In a similar vein, Pfeiffer et al. (1973), who studied the teaching job, found support for the content approach to multidimensional scaling.

While the content method of Ekman (1965) is one of several similarity methods which has found wide application and support, one may not wish to employ the method in its original form where bipolar factors are sought, or in the case where one is interested in obtaining the perceptual structure based on a single subject. For such applications, the Stone, Coles, and Lindem (1970) revision of the Ekman method may be employed. The Stone, Coles, and Lindem revision treats the similarity matrix given by the subject as raw data, intercorrelates the similarity estimates in the matrix, column by column, and then enters the resulting transformation, a correlational similarity matrix, for factor analysis. By this statistical technique, the many judgments of the paired columns are converted into single judgments possessing ratio scale properties. After all columns have been paired two at a time, a symmetrical matrix of true correlation coefficients numbering as many as the original raw similarity matrix results. Accordingly, meeting the strict assumptions of traditional factor analysis represents no problem for the user and, similarly, causes no restriction on the interpretation of the results.

Overview of Approach

To accomplish the goals of the total present program, a systematic set of steps was followed. These steps are presented in Figure 1-1. First, a set of military field messages was obtained. These messages were linguistically analyzed to yield kernels and operators, and these kernels and operators were grouped into a set of logical kernel classes and operator classes. Then a sample of messages representing each kernel and each operator class was subjected to a set of multidimensional scaling analyses to yield the dimensions (perceptual factors) involved. The methods, procedures, and results of these steps constitute the basis for Chapter II of the present report. The emergent dimensions were next woven into a taxonomy or classification scheme into which military field messages can be cast. Chapter III describes the taxonomy. Then the utility/reliability of the derived taxonomy was tested from both the information classification and the information extraction points of view (Chapter IV). Finally, the taxonomy was verified against the opinions of experienced intelligence analysts, and a computer oriented system for automatic processing of intelligence information was conceptualized (Chapter V).

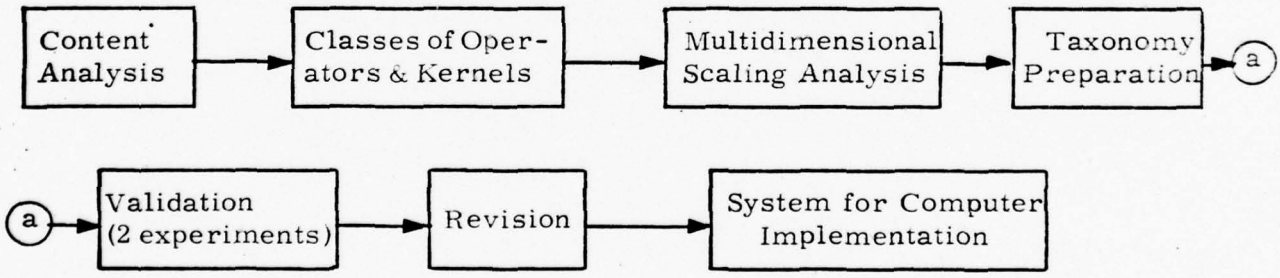


Figure 1-1. Program sequence.

CHAPTER II

MULTIDIMENSIONAL SCALING ANALYSIS

Message Set and Linguistic Analysis

In order to derive a battlefield oriented classification scheme, a realistic message set was required. A set of 110 messages was provided by the U. S. Army Research Institute for the Behavioral and Social Sciences. The messages were part of a scenario which concerned a worsening cold-war situation on the border between East and West Germany. For the most part, the messages contained two to four sentences each. All messages contained information about enemy activities. Accordingly, the classification scheme emerging from the present work may be limited to enemy activities. Examples of typical messages in the set of 110 are:

82. Imagery indicates the presence of a large storage area northeast of Werda

83. Observation of enemy activity along east side of Wiesse-Elster River indicates preparation for laying pontoons

Each message was then analyzed according to the theory of transformational grammar. Thus, corresponding to each sentence, a set of kernels and operators was derived in such a way that the original sentence was transformationally derived from that set. For example, a sentence taken from one message was: Friendly civilians report that two aggressor bns located in Schoneck as of 120900 Sep have been having difficulty in maintaining discipline within the town. A transformational decomposition of this sentence yields the kernel sentences: aggressor bns maintain discipline and aggressor bns are located in Schoneck. The operators are friendly civilians report; as of 120900 Sep; within the town; and have been having difficulty. According to the theory, each operator when applied to any of the sentences to which it is truly applicable has an effect of a specified type. For example, the first operator, friendly civilians report, has the effect of making a verb phrase complement of its operand, so that the new transformed sentence has friendly civilians as its subject and report as its verb. For the operators as of 120900 Sep and within the town, the sentences on which they act merely take the operators as adverbs. For the operator have been having difficulty, the result is to make the verb and object of the original sentence into a prepositional phrase, so that the verb of the transformed sentence is the verb of the operator.

In addition, there is an adjunct operator, two, which can be placed before the phrase aggressor bns. So far, forms friendly civilians report that two aggressor bns have been having difficulty in maintaining discipline within the town and aggressor bns are located in Schoneck as of 120900 Sep have been accounted for. The conjunction of these two forms, using the operator which followed by an operator which deletes redundancies (i.e., aggression bns) accounts for the final form of the message.

The performance of a transformational analysis of such material is not a cut and dried one. Since a sublanguage of English and not the entire language was involved, decisions were required at each step as to the correct way to analyze a sentence within that sublanguage. For example, in the entire language, the phrase aggressor troops must be analyzed as a transformed sentence like troops belong to aggressor, but in the sublanguage, the question has to be asked whether the phrase should be analyzed in such a fashion or remain unanalyzed. The latter alternative is the most reasonable one when the phrase friendly troops or its synonyms are not likely to appear in the language; even when such a phrase does appear, leaving aggressor troops unanalyzed might be worthwhile for simplicity's sake.

As a result of the transformational analysis, a set of kernels and operators sufficient to derive each of the sentences of the message set was obtained. The total set contained about 400 kernels and about 250 operators.

Classification (Content Analysis) of Kernels and Operators

The next step involved sorting the kernels and operators into content classes. The basis for the content analysis was that two kernels or operators were placed into the same class only if some meaningful classifier relationship existed between the two. For example, one class consisted of all operators which were of the adverbial form meaning "in the vicinity of a location"; one class of kernels consisted of verbs which were paraphrasable by the term "defensive." Some such linguistic justification was at the root of each of the classes obtained. The kernels were grouped into 25 classes. The operators were grouped into 23 classes. The 25 kernel classes are presented in Table 2-1. Table 2-2 presents the operator classes.

Table 2-1

Names of Kernel Classes

Civilian-military relations
Contents of documents
Actions of enemy
Weapons, military information
Sabotage, infiltration on both sides
Sources of and methods of obtaining information
Composition by name
Information about officers
Strategy, tactics, doctrine
"Is not known"
Markings and insignia
Replacements
Type, characterization of weapons, etc.
Compositions of weapons and of men
Actual and threatened contact
Sources of intelligence
Strength of units
Movement
Training, competence, etc.
Composition by number and type of equipment or by number of men
Morale and health
Characterizations of missions, roles, etc.
Analysis
Information about informers, individual soldiers, etc.
Location

Table 2-2

Names of Operator Classes

Aspectuals - training
Aspectuals - intention
Aspectuals - other
Adverbs - position - vicinity
Adverbs - position - other
Adverbs - position - unlocalized
Binaries
Adverbs - time
Adjuncts - on location adverbs
Adjuncts - enemy activities
Adjuncts - mechanical
Adjuncts - other
Significance raisers or reducers
Adverbs - force changers
Adverbs - other
Sentence operators - high reliability
low reliability
deserter
civilian
training, capability
intention
significance reducer
Unit membership

Stimulus Preparation for Multidimensional Scaling

The multidimensional scaling model requires that each member of the set to be factored be pair compared, relative to similarity, with every other member. These comparative judgments form the basic data set for the factor analysis which is the final step of the multidimensional scaling analytic procedure. To this end, one representative was selected from each class of operators and kernels and then one additional stimulus was selected from the 14 classes (7 operators and 7 kernels) containing the most stimuli. This procedure yielded a set of 62 stimuli. This final stimulus set is presented as Table 2-3.

Table 2-3

List of Military Intelligence Statements Employed in the Multidimensional Similarity Analysis

1. UNITS RECEIVE TROOP REPLACEMENTS
2. FORTIFICATIONS ARE PREPARED
3. SUPPLIES ARE RECEIVED
4. SABOT. TEAM INFILTRATES US LINES
5. TROOP UNIT MOVES SE
6. AGGRESSOR VEHICLES PASS THROUGH
7. HEAVY GUN EMPLACEMENTS PREPARED
8. DIVISION HAS NEW RIFLE
9. AGGR. RECONNAISSANCE TEAM SEEN
10. FIRE IS EXCHANGED
11. RANGE OF RIFLE IS INCREASED
12. PONTOON SECTIONS SUPPORT 60 TONS
13. POPULATION COOPERATES WITH AGGR.
14. 1F BN. IS IN SUPPORT OF 1F DIV.
15. AGGR. RAIL SYSTEM IS SABOTAGED
16. DOCUMENT ORIGINATES FROM 16 CAA
17. SPEED IS STRESSED
18. OFFICERS GRADUATED WAR COLLEGE
19. STRENGTH IS 83% OF TO&E
20. COURT MARTIALS ARE DISCLOSED
21. UNITS INFECTED WITH INFLUENZA
22. UNIT IS LOCATED UR037847
23. MEN RESPECTED COMMANDER
24. UNIT IS AMPHIBIOUS
25. PATROL CONSISTS OF 20 MEN
26. TRAIN IS LOADED WITH SUPPLIES
27. AGGR. COMMUN. DISCIPLINE POOR
28. CG ENTERS MANEUVER AREA
29. PONTOONS ARE LAYED
30. MINE FIELDS ARE PREPARED
31. UNIT IS ENGINEERING UNIT
32. UNIT IS MECH. UNIT
33. DESERTER ENTERED FRIENDLY LINES
34. AGGR. UNITS ARE WELL TRAINED
35. ...IS CAPABLE OF...
36. MISSION IS TO...
37. ...IN VICINITY OF TOWN
38. ...IN TOWN
39. ...THROUGHOUT AGGRESSOR AREA
40. ...OCCURRED
41. ...AT THIS TIME
42. ...IMMEDIATELY
43. ...IN THE FUTURE
44. ...LARGE VOLUME OF...
45. ...SMALL NUMBER OF...
46. AGGR. DOCUMENT INDICATES...
47. RADIO BROADCASTS INDICATE...
48. US RADAR SECTION REPORTS...
49. DESERTER CLAIMS THAT
50. CIVILIAN INDICATES...
51. ...DID NOT OCCUR
52. AGGRESSOR EXPECTS THAT...
53. ...IS AERIAL
54. 16CAA AUTHORIZES...
55. ENEMY DIV. IS ALERTED TO...
56. ...INCREASES
57. ...IS RAPID
58. ...DECREASES
59. COMMANDER STATES...
60. ...IS LOGISTICAL
61. ...IS PREPARING TO...
62. ...NEAR BORDER

Analysts and Procedure

To obtain similarity estimates between pairs of military intelligence statements, as required by the Stone et al. (1970) method of multidimensional similarity analysis, four retired military personnel with extensive operational experience were employed as analysts. The military and intelligence experience of these persons is presented in Table 2-4.

Table 2-4

Military Background and Experience of Analysts

Subject	Rank*	Years of Military Service by Type		
		Military Intelligence	Other	Total
Mc	Lt. Col.	1	21	22
M	Capt.	4	0	4
K	Capt.	3	1	4
S	Lt. Col.	13	3	16

* Rank at discharge

Each analyst was briefed in a separate data collection session by an Applied Psychological Services' staff member. The briefing included: (1) an explanation of the purposes and background of the study, (2) a description of military intelligence message list and the basis for its development, and (3) an explanation of the manner in which the similarity data would be collected. A scale of similarity values was distributed to each judge; each judge used the scale in making magnitude estimations of the degree of similarity between all pairs of messages on a 0 to 100 per cent scale. Since a large number of comparisons [(62 x 61)/2] was involved, each analyst responded to one-third the required estimations on one of three separate data collection sessions. Each data collection session involved about three hours work. The analysts were paid for their participation.

For control purposes, the stimulus pairs were presented to one-half of the analysts in a random order which was the reverse of the random order of the stimuli for the other analysts. To check on intra-analyst reliability, ten pairs of stimuli were presented to each of the analysts twice--once in the form in which the stimuli to be judged were in the order A-B and the other time with the stimuli in the order B-A.

Results

The reliability check stimuli were interspersed randomly throughout the stimulus set. The product moment correlation was calculated between the paired data sets for each of the four analysts. The product moment correlations so derived were: .68, .98, .31, and .64. The mean (employing z' transformation) reliability is .77*. These data suggest reasonable subject consistency and minimum effect due to the directionality of the comparisons involved.

Factor Analysis

The similarity estimates given by the four analysts resulted in four separate similarity matrices and one based on the average of the four. Each matrix was 62 x 62 in dimension. In accordance with the Stone, Coles, and Lindem (1970) method, the columns of each raw similarity matrix were intercorrelated. The result of this step was a product moment correlation similarity matrix containing the same number of elements as the original. Then, each analyst's correlation similarity matrix, as well as the averaged data matrix, was subjected to factor analysis by the principal components method with orthogonal rotation in accordance with the Kaiser developed varimax criterion (Dixon, 1973; BMD 03M, Revised 9-1-65). Communalities were estimated at 1.00. Factor extraction was stopped after the 16th factor. The cumulative percentage of unrotated variance accounted for in each analysis is presented in Table 2-5. Table 2-5 indicates that for the individual analyses, 78 to 82 per cent of the total variance was accounted for by the 16 factor solution. For the averaged matrix, 86 per cent of the variance was accounted for.

* When this value is corrected by the Spearman-Brown technique, the overall reliability becomes .93.

Table 2-5

Cumulative Unrotated Percentage of Variance Accounted for by Each Factor Analysis

Factor	Average	Subject			
		Mc	M	K	S
1	20	18	16	20	23
2	33	31	28	29	37
3	44	40	36	35	46
4	51	47	43	41	52
5	57	53	48	46	57
6	61	58	53	51	61
7	66	62	56	55	64
8	69	66	60	59	67
9	73	69	63	62	71
10	75	71	66	64	73
11	78	73	69	67	75
12	80	75	72	70	77
13	81	77	74	72	78
14	83	79	76	74	80
15	85	80	78	76	82
16	86	82	79	78	83

Agreement Across Analyses

The next step involved assessing the agreement or "congruency" of the factors across the individual factor analysis for each of the four analysts. If the four separate analyses yielded factors which are common, some generality can be attributed to the factors and confidence can be placed in the factor analysis which integrated the data of the four subjects. If little or no agreement among the perceptual structures of the four analysts is found, then a structure which possesses general utility (at least for the message set involved) cannot be held to have been accomplished by the integrated analysis.

Accordingly, each rotated factor solution was compared with every other rotated solution, making a total of six comparisons among four analysts. The comparison was accomplished in accordance with the procedures developed by Kaiser, Hunka, and Bianchini (1971). This method seeks the degree of agreement in factorial structures by a cosine rotation method. In each of the studies being compared two at a time, the geometric configuration of the factor vectors, in the space where cosines of the angles between vectors signify correlations, are considered. Accordingly, the cosine of the angle between any two vectors may be interpreted as the correlation between the two variables (factors) represented.

For the present investigation, the rotated factorial structures of the four analysts were compared taking them two at a time, making a total of six comparisons. Since each factor analysis resulted in 16 factors, a 16 x 16 matrix of 256 cosines representing the relationship of all rotated factors of one analyst with all rotated factors of the comparison analyst resulted. The cosine matrices for all comparisons are presented in Tables 2-6 through 2-11.

Since the 16 x 16 cosine matrices constituted excessive detail, an appropriate logic and technique summarizing the cosines across factors was sought. Butzkamin and Pfeiffer (1973) were faced with a similar problem in their cross-cultural study of the perceived structure of teaching. To assess the average degree of agreement of factors across two samples based upon different individuals, the following index of agreement was employed:

$$Ai = \left[1 - \left(\frac{\sum_{i=1} C_{si}^2}{N} \Big/ C_L^2\right)\right]^2$$

where:

Ai = index of agreement

C_L = the largest cosine in any column/row of the cosine matrix indicating which factor from one study shows most agreement with the corresponding factor from the other study. In the event of a controversy over the identification of the column/row intersection, an additional cosine is identified.

C_{si} = the remaining smaller cosines from the same column/row as above, representing factors having smaller agreements.

N = the total number of possible non-corresponding factors in the columns/row of the matrix; 30 in the present case where matrix size is 16 x 16.

To calculate the index, first the row/column intersect possessing the highest value for each column was identified. In the case of a 16 x 16 matrix (without controversies) there were 16 such intersects. Each intersect indicated the two factors possessing the greatest congruency. Then the formula was applied. The advantage of this index over other similar indices, such as the Fisher z' technique of Kaiser, Hunka, and Bianchini (1971), is that the present index considers both agreement and disagreement. Other indices only consider agreement. The value of Ai can range from zero (no agreement) to one (perfect agreement).

Table 2-6

Cosines Among All Pairs of 16 Intelligence Communications Factors for Analysts M and Mc

Factors for Subject M	Factors for Subject Mc															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
FACTORS FOR SUBJECT M 1	-0.24	-0.18	-0.03	0.12	-0.34	-0.08	-0.19	0.10	0.07	-0.00	-0.03	-0.01	-0.83	-0.04	-0.14	-0.06
2	-0.14	0.06	0.45	0.01	-0.14	0.10	0.71	0.28	0.07	0.15	0.37	-0.02	-0.07	0.00	-0.11	0.07
3	0.09	-0.09	-0.06	-0.08	0.12	-0.92	0.20	0.13	0.07	-0.10	-0.01	0.15	-0.03	-0.05	0.22	-0.03
4	-0.21	0.82	-0.02	0.04	0.05	-0.23	-0.12	-0.05	0.12	0.26	-0.04	-0.26	-0.05	0.21	-0.13	0.03
5	0.32	0.16	0.06	0.46	-0.07	0.08	-0.14	0.28	0.52	0.05	-0.05	0.11	0.08	-0.48	0.15	-0.01
6	-0.03	-0.19	-0.12	0.49	0.49	0.01	0.19	-0.02	-0.11	-0.19	0.07	-0.59	-0.12	-0.03	-0.02	-0.14
7	-0.19	0.06	0.23	-0.06	0.17	0.02	-0.11	-0.61	0.11	-0.00	0.48	0.11	-0.15	-0.27	0.36	-0.13
8	0.00	0.07	-0.64	-0.06	-0.12	0.05	0.28	-0.07	0.01	0.28	0.18	0.12	0.04	-0.21	-0.19	-0.53
9	0.26	0.33	0.12	-0.06	0.02	0.02	-0.09	0.24	-0.73	-0.10	0.07	0.06	-0.17	-0.36	0.11	-0.11
10	0.09	-0.25	0.16	-0.04	0.34	-0.05	-0.20	0.13	-0.09	0.85	-0.04	-0.00	-0.07	0.03	0.03	-0.01
11	-0.10	0.00	0.04	0.64	-0.01	-0.08	-0.06	-0.08	-0.25	0.01	0.16	0.55	0.11	0.36	-0.14	-0.06
12	-0.05	-0.16	0.14	0.19	-0.52	-0.28	-0.09	-0.28	-0.22	0.18	0.05	-0.36	0.33	-0.32	-0.24	0.12
13	0.02	0.08	0.29	0.08	0.04	0.02	0.34	-0.36	-0.04	0.06	-0.73	0.14	-0.11	-0.12	-0.02	-0.27
14	-0.57	-0.03	-0.05	0.10	-0.21	0.08	-0.02	0.29	-0.12	0.07	-0.16	-0.10	0.21	0.02	0.62	-0.21
15	-0.57	-0.04	0.06	-0.11	0.36	-0.03	-0.13	0.19	0.02	-0.14	-0.07	0.20	0.16	-0.43	-0.45	0.05
16	0.11	-0.06	0.40	-0.12	-0.05	-0.09	-0.27	0.16	0.13	-0.16	0.10	-0.14	0.14	0.20	-0.18	-0.73

Table 2-7

Cosines Among All Pairs of 16 Intelligence Communications Factors for Analysts Mc and K

Factors for Subject Mc	Factors for Subject K															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	-0.07	0.15	0.07	0.74	-0.01	0.16	-0.17	0.10	0.02	0.10	-0.06	0.03	0.40	0.02	0.03	0.42
2	-0.20	0.85	-0.06	0.01	-0.06	-0.07	0.19	0.11	-0.00	0.06	0.23	0.29	-0.20	-0.11	0.03	-0.11
3	0.54	0.26	0.16	0.04	0.56	0.18	-0.13	-0.47	-0.04	0.17	-0.28	0.04	-0.08	0.18	-0.23	-0.16
4	-0.48	0.01	-0.07	-0.07	0.28	0.11	0.15	0.01	-0.16	0.05	-0.71	0.01	-0.04	-0.30	0.11	0.01
5	-0.05	0.08	-0.21	-0.25	-0.06	0.21	-0.17	-0.58	0.09	0.12	0.07	0.11	0.13	-0.29	-0.48	0.32
6	0.20	-0.05	0.01	-0.11	-0.14	0.82	0.17	0.07	-0.08	-0.01	0.10	0.12	-0.16	-0.03	0.32	0.06
7	0.43	0.17	-0.02	-0.09	0.05	-0.30	0.40	-0.02	0.03	-0.22	-0.08	-0.28	-0.15	-0.24	0.19	0.57
8	0.09	0.29	0.21	-0.26	-0.39	0.12	0.10	0.11	0.38	-0.09	-0.33	-0.30	0.46	0.08	-0.11	-0.19
9	-0.13	0.02	-0.12	-0.07	0.35	0.15	0.02	0.15	-0.09	-0.80	0.14	0.11	0.26	0.12	-0.22	0.04
10	-0.21	0.04	0.74	-0.04	0.11	-0.02	-0.10	-0.45	0.16	-0.20	0.05	0.12	-0.09	0.08	0.28	0.09
11	0.05	0.11	-0.46	0.15	-0.18	0.01	-0.37	-0.28	0.26	-0.30	-0.32	0.04	-0.29	0.31	0.27	-0.02
12	0.20	-0.02	0.07	-0.31	0.12	0.	-0.15	0.48	0.40	0.19	-0.01	0.06	-0.28	0.29	-0.16	0.47
13	-0.03	0.00	0.13	0.04	-0.35	-0.04	0.26	-0.14	-0.51	-0.04	-0.22	0.13	-0.13	0.53	-0.34	0.16
14	0.29	-0.16	0.10	-0.06	-0.21	-0.20	-0.07	0.19	0.07	-0.11	-0.24	0.77	0.10	-0.27	-0.00	0.01
15	-0.02	-0.12	0.18	0.41	-0.13	0.12	0.13	0.06	0.28	-0.22	-0.04	-0.13	-0.49	-0.30	-0.46	-0.18
16	0.02	0.15	0.18	-0.13	-0.22	0.05	-0.64	0.22	-0.47	-0.15	-0.05	-0.25	-0.14	-0.27	-0.02	0.08

Table 2-8

Cosines Among All Pairs of 16 Intelligence Communications Factors for Analysts M and K

FACTORS FOR SUBJECT M	Factors by Subject K															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	-0.10	-0.18	-0.15	-0.28	0.42	-0.08	0.02	0.29	0.27	0.07	0.33	-0.47	0.06	-0.23	0.19	-0.32
2	0.59	0.37	-0.07	-0.11	0.36	-0.08	-0.18	-0.12	-0.18	-0.29	-0.10	-0.23	-0.23	0.22	0.05	-0.00
3	0.04	0.16	0.05	0.19	-0.04	-0.60	-0.32	-0.05	0.51	-0.10	-0.21	0.05	0.37	-0.00	-0.02	-0.08
4	-0.32	0.62	0.05	-0.42	0.09	-0.25	0.35	-0.00	0.04	-0.06	0.05	0.15	-0.14	-0.24	-0.13	0.22
5	-0.55	0.29	-0.06	0.36	0.33	0.24	-0.22	0.05	-0.23	-0.38	0.05	-0.01	0.25	0.08	-0.13	-0.12
6	0.10	-0.13	-0.33	-0.10	0.14	0.14	0.07	-0.69	-0.07	-0.02	-0.21	0.06	0.28	-0.45	0.21	-0.11
7	-0.25	-0.27	-0.46	0.30	0.15	-0.31	0.04	-0.14	0.10	-0.02	-0.11	-0.07	-0.53	-0.01	-0.10	0.28
8	-0.32	-0.06	0.04	-0.28	0.05	-0.12	-0.06	-0.30	-0.02	0.09	0.11	0.02	0.08	0.63	0.51	0.14
9	0.00	0.47	-0.20	0.33	-0.03	0.06	-0.13	-0.05	-0.08	0.72	0.06	-0.23	-0.03	-0.01	0.11	-0.03
10	-0.11	-0.08	0.62	0.07	0.13	0.11	0.03	-0.42	0.16	0.07	-0.06	-0.50	-0.08	-0.04	-0.28	0.14
11	-0.20	-0.05	0.06	-0.33	0.18	0.20	-0.37	0.23	0.03	0.22	-0.71	0.09	-0.16	-0.04	0.02	-0.04
12	0.09	0.06	0.15	0.23	0.41	0.16	0.34	-0.08	0.38	0.12	0.00	0.46	-0.21	0.19	0.04	-0.37
13	0.18	-0.22	0.33	0.18	0.42	-0.43	0.09	0.14	-0.52	0.20	-0.04	0.16	0.15	-0.18	0.14	0.14
14	-0.09	0.01	-0.09	0.03	-0.18	-0.21	0.50	0.04	-0.23	-0.05	-0.43	-0.33	0.07	0.25	-0.07	-0.47
15	-0.10	-0.08	0.04	-0.21	-0.07	-0.24	-0.38	-0.26	-0.25	0.09	0.28	0.18	-0.28	0.00	-0.35	-0.51
16	0.12	-0.14	-0.29	-0.21	0.28	0.04	0.08	-0.01	0.03	0.25	0.03	0.05	0.40	0.33	-0.61	0.21

Table 2-9

Cosines Among All Pairs of 16 Intelligence Communications Factors for Analysts Mc and S

Factors by Subject S

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
FACTORS FOR SUBJECT Mc	1	0.10	0.56	-0.07	0.03	-0.40	-0.04	-0.37	-0.18	0.20	-0.38	0.28	0.09	0.05	0.09	-0.24	-0.08
	2	0.04	0.08	-0.16	0.72	0.10	-0.23	-0.20	0.22	-0.10	0.32	0.20	-0.33	0.05	-0.01	-0.13	0.05
	3	0.69	0.26	0.02	0.05	-0.07	0.08	0.38	0.00	0.14	0.12	-0.24	-0.22	-0.03	0.24	0.12	-0.31
	4	0.25	0.36	0.08	-0.03	0.27	-0.14	0.11	0.22	0.00	0.06	-0.17	0.40	-0.22	-0.45	-0.31	0.32
	5	0.01	-0.06	0.08	-0.31	-0.45	-0.20	0.28	0.46	0.00	0.19	0.23	-0.11	0.21	0.20	-0.26	0.32
	6	-0.16	0.05	0.10	-0.03	-0.11	0.03	0.03	0.04	0.01	0.08	0.17	-0.13	-0.92	0.21	0.00	0.01
	7	0.08	0.09	-0.04	0.15	0.12	-0.04	0.22	0.02	0.24	-0.12	0.48	0.22	0.03	0.02	0.65	0.34
	8	-0.06	-0.19	-0.46	0.27	-0.14	0.11	0.11	0.24	0.26	-0.46	-0.44	-0.02	-0.10	0.12	-0.07	0.28
	9	0.05	0.05	-0.14	-0.00	0.08	0.89	-0.02	0.06	0.06	0.23	0.23	-0.01	0.06	-0.02	-0.21	0.14
	10	0.39	-0.29	0.04	-0.12	-0.21	0.01	-0.16	-0.14	-0.04	-0.19	0.12	-0.47	-0.12	-0.56	0.07	0.22
	11	-0.37	0.54	-0.30	-0.17	-0.15	0.01	0.14	-0.12	-0.16	0.20	-0.26	-0.33	0.05	-0.22	0.29	0.15
	12	0.01	-0.10	-0.53	-0.33	0.31	-0.26	0.02	-0.15	0.51	0.19	0.19	-0.12	-0.04	-0.04	-0.21	-0.14
	13	-0.32	0.14	0.48	0.11	0.24	0.06	0.30	0.10	0.40	-0.29	0.05	-0.39	0.10	-0.17	-0.16	-0.13
	14	0.04	0.11	-0.03	-0.27	0.16	0.05	-0.45	0.70	0.05	-0.10	-0.04	-0.09	-0.01	-0.04	0.31	-0.26
	15	0.10	0.12	-0.19	-0.15	0.43	-0.03	0.17	-0.00	-0.56	-0.44	0.21	-0.22	0.01	0.25	-0.16	0.09
	16	0.10	-0.08	-0.28	0.13	-0.26	0.04	0.39	0.19	-0.21	-0.11	0.26	0.21	-0.06	-0.41	-0.03	-0.54

Table 2-10

Cosines Among All Pairs of 16 Intelligence Communications Factors for Analysts M and S

FACTORS FOR SUBJECT M	Factors by Subject S															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.18	-0.44	-0.21	0.02	0.17	0.15	-0.30	-0.31	-0.04	0.30	-0.18	0.56	0.01	-0.15	0.13	0.03
2	0.28	0.22	-0.16	0.29	0.15	0.20	0.28	0.04	0.22	-0.06	0.00	0.00	-0.16	0.16	0.67	0.26
3	0.07	-0.06	-0.45	-0.01	-0.04	0.01	-0.04	0.19	-0.08	-0.13	-0.07	-0.06	0.80	0.18	-0.03	0.22
4	0.19	-0.06	0.15	0.49	0.09	-0.24	-0.18	0.24	0.22	0.44	0.25	-0.23	0.20	-0.37	0.00	-0.05
5	0.21	0.24	-0.03	0.26	0.15	0.44	-0.14	0.05	0.23	0.00	-0.10	0.05	-0.15	0.16	-0.64	0.27
6	0.08	0.20	0.46	-0.26	-0.22	-0.23	-0.00	0.13	-0.02	0.21	-0.09	0.32	0.10	0.00	0.05	0.62
7	0.06	0.16	0.01	0.16	0.11	0.22	0.34	-0.20	-0.75	0.20	0.29	-0.01	0.07	-0.16	-0.09	0.11
8	-0.11	0.05	0.24	0.07	0.13	0.20	-0.65	-0.12	-0.28	-0.23	-0.17	-0.39	0.01	-0.07	0.26	0.21
9	0.14	-0.06	-0.16	0.53	-0.41	-0.44	-0.08	0.00	-0.30	-0.24	-0.21	0.12	-0.25	0.12	-0.08	0.07
10	0.50	-0.37	0.07	-0.19	-0.00	-0.07	0.16	-0.25	0.12	-0.43	0.19	-0.21	-0.03	-0.33	-0.13	0.26
11	0.17	0.32	-0.32	-0.24	0.02	0.01	-0.00	0.37	-0.12	-0.03	-0.40	-0.00	-0.15	-0.61	0.01	-0.09
12	-0.07	0.27	0.06	0.09	0.65	-0.49	0.11	-0.33	0.05	-0.15	-0.25	0.08	0.15	-0.01	-0.10	-0.02
13	0.62	0.11	0.39	-0.07	-0.02	0.08	-0.05	0.08	-0.15	-0.05	-0.16	0.09	0.20	0.27	0.05	-0.51
14	-0.12	-0.43	0.19	0.07	0.43	0.02	0.12	0.64	-0.20	-0.22	-0.00	0.17	-0.12	0.07	-0.02	0.08
15	-0.02	-0.32	0.10	0.03	-0.04	0.06	0.35	-0.08	-0.01	0.36	-0.63	-0.45	-0.00	0.09	-0.05	0.09
16	-0.28	0.00	0.31	0.34	-0.23	0.31	0.22	-0.09	0.17	-0.32	-0.20	0.28	0.31	-0.39	0.03	-0.11

Table 2-11

Cosines Among All Pairs of 16 Intelligence Communications Factors for Analysts K and S

Factors by Subject S

FACTORS FOR SUBJECT K	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	0.02	0.15	-0.05	0.07	-0.13	0.05	0.13	-0.19	0.12	-0.19	-0.24	0.02	-0.00	0.57	0.67	-0.07
2	0.19	0.26	-0.25	0.72	0.08	-0.23	-0.01	0.15	0.08	0.13	0.13	-0.18	0.27	0.19	-0.16	0.13
3	0.51	-0.42	-0.11	-0.08	0.11	-0.05	-0.21	0.10	0.26	-0.44	0.09	-0.37	-0.24	0.11	-0.06	-0.01
4	0.03	0.15	0.16	0.27	-0.48	0.07	-0.06	0.26	0.18	-0.21	0.39	0.38	-0.37	-0.09	0.03	-0.21
5	0.74	0.17	-0.10	-0.05	-0.17	0.20	0.18	-0.19	-0.01	0.38	-0.20	0.13	-0.12	-0.27	0.07	0.08
6	-0.04	0.38	0.05	-0.19	-0.01	0.07	-0.03	-0.12	-0.05	0.11	-0.01	-0.10	-0.48	0.51	-0.50	0.17
7	0.18	0.05	0.59	0.25	0.14	0.30	-0.33	0.05	-0.52	-0.04	-0.04	-0.18	0.03	0.08	0.10	-0.16
8	0.02	0.10	-0.45	0.06	0.39	0.05	-0.25	-0.11	-0.12	-0.03	-0.12	0.30	-0.15	-0.01	-0.08	-0.64
9	0.09	-0.08	-0.13	-0.35	-0.23	0.15	-0.25	0.36	0.02	0.45	0.28	-0.08	0.33	0.31	0.06	-0.25
10	0.15	-0.13	0.37	-0.04	-0.18	-0.71	-0.12	-0.11	0.08	0.15	-0.30	0.18	0.07	0.12	-0.16	-0.25
11	-0.12	-0.50	-0.17	0.28	-0.38	0.11	0.05	-0.55	-0.23	0.15	0.18	-0.15	-0.07	0.08	-0.12	-0.12
12	-0.14	0.16	-0.22	0.01	-0.41	-0.08	0.08	0.33	-0.24	-0.11	-0.43	-0.49	-0.13	-0.21	-0.02	-0.24
13	0.00	0.19	-0.12	-0.07	-0.34	0.21	-0.53	-0.26	0.13	-0.33	-0.22	0.13	0.40	-0.05	-0.22	0.19
14	-0.16	0.13	0.27	0.10	0.10	0.23	-0.03	-0.22	0.67	0.25	-0.08	-0.36	-0.01	-0.18	-0.02	-0.31
15	-0.07	0.23	-0.12	-0.10	-0.00	-0.36	-0.49	-0.21	-0.09	0.18	0.28	-0.22	-0.27	-0.22	0.40	0.17
16	0.17	0.36	0.07	-0.25	-0.03	-0.15	0.34	-0.28	-0.15	-0.33	0.43	-0.20	0.30	-0.05	-0.08	-0.32

Since a number of controversies were involved in row/column intersect selection, the number of comparisons involved in each case ranged from 17 to 21. The resultant indices of agreement are presented in Tables 2-12 through 2-17. The individual values presented in Tables 2-12 through 2-17 are all reasonably high.*

To obtain the average agreement of all factors of each analysis with all factors of each other analysis, the indices of agreement (Ai) were averaged across the total number of potentially corresponding factors, i.e., highest Ai values. The results are presented in Table 2-18. The average indices of agreement presented in Table 2-18 ranged from .71 to .83. These data suggest strong agreement among the factorial structures of the data yielded by the four separate analysts.

When the analysts are ordered in terms of their average degree of agreement with each other (Table 2-18), the results indicate the following order from most to least interjudge agreement:

$$Mc > M > K > S$$

Accordingly, analyst Mc (the analyst with the most general military and least military intelligence experience) may be viewed as having a perception of intelligence communications most common, or most central, to the other experts. However, it is noted that the differences among the analysts are small. The reason for this indication (if, in fact, true) is not known. It may be due to the common Army background of the analysts.

---

* Additionally z' agreement coefficients were calculated using the method suggested by Kaiser, Hunka, and Bianchini (1971). Only trivial differences were noted between the results indicated by the two approaches.

Table 2-12

Agreement (Ai) Among Potentially Congruent **Factors** Produced by Analysts M and Mc

Factors of M	Mc	Ai
1	13	0.94
2	7	0.87
3	6	0.97
4	2	0.94
5	9	0.67
6	12	0.78
7	8	0.78
8	3	0.82
9	9	0.89
10	10	0.95
11	4	0.82
12	5	0.67
13	11	0.89
14	15	0.80
15	1	0.73
16	16	0.88
5	14	0.60
Mean		0.82

Table 2-13

Agreement (Ai) Among Potentially Congruent Factors Produced by Analysts Mc and K

Factors of Mc	K	Ai
1	4	0.89
2	2	0.95
3	5	0.74
4	11	0.86
5	8	0.75
6	6	0.94
7	16	0.74
8	13	0.56
9	10	0.92
10	3	0.89
11	3	0.57
12	8	0.60
13	14	0.69
14	12	0.91
15	13	0.63
16	7	0.82
3	1	0.71
13	9	0.66
5	15	0.61
Mean		0.76

Table 2-14

Agreement (Ai) Among Potentially Congruent Factors Produced by Analysts M and K

Factors of M	Factors of K	Ai
1	12	0.58
2	1	0.76
3	6	0.78
4	2	0.78
5	1	0.70
6	8	0.85
7	13	0.69
8	14	0.81
9	10	0.89
10	3	0.79
11	11	0.87
12	12	0.57
13	9	0.67
14	7	0.65
15	16	0.67
16	15	0.79
4	4	0.46
13	5	0.48
Mean		0.71

Table 2-15

Agreement (Ai) Among Potentially Congruent Factors Produced by Analysts Mc and S

Factors of Mc	S	Ai
1	2	0.73
2	4	0.88
3	1	0.86
4	14	0.55
5	8	0.57
6	13	0.98
7	15	0.83
8	10	0.57
9	6	0.96
10	14	0.73
11	2	0.70
12	3	0.69
13	3	0.60
14	8	0.87
15	9	0.73
16	16	0.70
5	5	0.53
14	7	0.55
7	11	0.61
10	12	0.58
Mean		0.71

Table 2-16

Agreement (Ai) Among Potentially Congruent Factors Produced by Analysts M and S

Factors of M	Factors of S	Ai
1	12	0.73
2	15	0.84
3	13	0.92
4	4	0.62
5	15	0.82
6	16	0.80
7	9	0.90
8	7	0.83
9	4	0.69
10	1	0.64
11	14	0.78
12	5	0.83
13	1	0.80
14	8	0.82
15	11	0.81
16	14	0.39
1	2	0.53
6	3	0.56
12	6	0.62
4	10	0.52
Mean		0.72

Table 2-17

Agreement (Ai) Among Potentially Congruent Factors Produced by Analysts K and S

Factors of K	Factors of S	Ai
1	15	0.84
2	4	0.88
3	1	0.66
4	5	0.61
5	1	0.89
6	14	0.67
7	3	0.77
8	16	0.81
9	10	0.54
10	6	0.87
11	8	0.72
12	12	0.63
13	7	0.69
14	9	0.84
15	7	0.62
16	11	0.51
11	2	0.63
6	13	0.60
Mean		0.71

Table 2-18

Average Agreement Among Four Factorial Structures

Subject	Mc	M	K	S
Mc	-	.823	.760	.712
M	.823	-	.711	.722
K	.760	.711	-	.710
S	.712	.722	.710	-
ΣA_i	2.29	2.26	2.18	2.16

Cluster Analysis

In an attempt to examine the issue of factorial agreement from a slightly different point of view, a hierarchical cluster analysis was performed (Johnson, 1967). Specifically, the averaged agreement indices (Table 2-18) were cluster analyzed to determine whether or not the (average) agreement among the subjects formed more than one cluster. A single cluster emerged. A spacial summary of the analysis is presented as Table 2-19. Again, the strength of the agreement suggesting a single cluster is strong. The strength for cluster entry ranged from .71 to .82. This finding suggests support for a composite 16 factor solution.

Table 2-19

Cluster Analysis Results Showing a Single Structure Among Agreement Indices for Four Analysts

Subjects				
Mc	M	S	K	Strength
XXX	XXX			.823
XXX	XXX	XXX		.712
XXX	XXX	XXX	XXX	.710

Final Factor Analysis

The agreement results derived from the data emerging from the cosine analysis along with the results of the hierarchical cluster analysis supported contentions favoring a high degree of agreement among the four analysts in the perception of army field messages. Accordingly, one final structure was sought which best reflected the combined perceptions of the four analysts. Accordingly, the raw similarity matrices of the four analysts were averaged, correlational similarities computed, and a single matrix was submitted to factor analysis and rotation. Sixteen factors were again extracted. The final rotated factor loadings are presented as Appendix A to this report. The eigenvalue for each factor was:

Factor	Eigenvalue	Factor	Eigenvalue
I	12.3	IX	2.0
II	8.1	X	1.7
III	6.7	XI	1.5
IV	4.2	XII	1.2
V	3.7	XIII	1.1
VI	3.0	XIV	1.0
VII	2.5	XV	1.0
VIII	2.4	XVI	0.8

We note that the eigenvalue of the last factor is less than the usual arbitrary cut point of one. Some degree of overfactoring seems defensible in the present case where comprehensiveness is sought. Factors can be eliminated if they seem unwarranted as the result of use tests. However, new empirical factors cannot be easily developed.

Description of Each Factor

Each of the 16 factors isolated in the final, overall factor analysis is described in this section. Each factor is also named on the basis of stimulus items possessing heavy loadings on the factor.

The stimulus items possessing heavy loadings on factor I are presented in Table 2-20. The stimulus items loading heavily on this factor all seem to be concerned with movement and potential for movement. Accordingly, this factor was named Movement and Capability for Movement.

Table 2-20

Stimulus Items Loading Heaviest on Factor I--
Movement and Capability for Movement

Stimulus	Description	Loading
12	pontoon sections support 60 tons	-.90
29	pontoons are laid	-.84
17	speed is stressed	-.74
32	unit is mech unit	-.69
24	unit is amphibious	-.67
57	...is rapid	-.63
5	troop unit moves SE	-.59
6	aggressor vehicles pass through	-.56

A number of items concerned with numerical size, trends, and data sources loaded heavily on factor 2. These stimulus items are presented in Table 2-21. Factor 2 was named Numbers, Trends, and Source in view of such stimulus items as: "small number of," "decreases," and "US radar section reports." It is not clear why sources fell on this factor. Possibly, this construct should be separated out into a separate category.

Table 2-21

Stimulus Items Loading Heaviest on Factor II--
Numbers, Trends, and Source

Stimulus	Description	Loading
45	...small number of...	-.77
58	...decreases	-.76
44	...large volume of...	-.67
48	US radar section reports...	-.60
51	...did not occur	-.59
40	...occurred	-.54
53	...is aerial	-.35

The third factor was loaded on items describing the character of various operations, e.g., "sabotage team infiltrates US lines," "aggr. reconnaissance team seen." The stimulus items loading heaviest on this factor are presented in Table 2-22. The factor was named Type of Activity.

Table 2-22

Stimulus Items Loading Heaviest in Factor III--
Type of Activity

Stimulus	Description	Loading
9	aggr. reconnaissance team seen	-.88
25	patrol consists of 20 men	-.87
4	sabotage team infiltrates US lines	-.81
10	fire is exchanged	-.77

Factor IV seemed to reflect the state of readiness of the hostile forces as a function of their training, motivation/morale, and attitudes. This factor was named Training, Morale, and Attitudes. Table 2-23 presents the items loading heaviest on this factor.

Table 2-23

Stimulus Items Loading Heaviest on Factor IV--Training, Morale, and Attitudes

Stimulus	Description	Loading
18	officers graduated war college	-.85
23	men respected commander	-.82
34	aggr. units are well trained	-.72
20	court martials are disclosed	-.42

Factor V contained a somewhat mixed set of heavily loaded stimulus items. On the one hand, the factor alluded to communications, while on the other hand, it reflected railroad oriented information. Since railroad information would also group with factor I (Movement and Capability for Movement), factor V was named Communications. The stimulus items loading heavily on factor V are presented in Table 2-24.

Table 2-24

Stimulus Items Loading Heaviest on Factor V--Communications

Stimulus	Description	Loading
27	aggr. communication discipline poor	-.87
15	aggr. rail system is sabotaged	-.72
26	train is loaded with supplies	-.35

Factor VI was uniquely concerned with armament preparation and was named accordingly. The items loading heavily on this factor are presented in Table 2-25.

Table 2-25

Stimulus Items Loading Heaviest on Factor VI--Armament Preparation

Stimulus	Description	Loading
2	fortifications are prepared	-.87
7	heavy gun emplacements prepared	-.85
30	mine fields are prepared	-.76

The stimulus items loading heavily on factor VII were concerned with rifles. To generalize the factor, factor VII was named Small Arms Capability. The items loading heaviest on factor VII, Small Arms Capability, are presented in Table 2-26.

Table 2-26

Stimulus Items Loading Heaviest on Factor VII--Small Arms Capability

Stimulus	Description	Loading
8	division has new rifle	.93
11	range of rifle is increased	.89

Any descriptive scheme that aims to classify military field information would logically be expected to contain a category relative to the striking power and position of various units. Factor VIII was concerned with this type of information. The items loading heaviest on this factor are presented in Table 2-27, and the factor was named Strength and Location of Units to reflect the content of the heavily loaded stimulus items.

Table 2-27

Stimulus Items Loading Heaviest on Factor VIII--
Strength and Location of Units

Stimulus	Description	Loading
19	strength is 83% of TO&E	-.78
22	unit is located UR037847	-.73
21	units infected with influenza	-.66
1	units receive troop replacements	-.50
14	1F bn. is in support of 1F div.	-.46
37	...in vicinity of town	-.32

The stimulus items loading heaviest on factor IX are presented in Table 2-28. The logical thread which seems to weave these items together is intent and anticipation. Accordingly, the factor was termed Plans and Expectations.

Table 2-28

Stimulus Items Loading Heaviest on Factor IX--
Plans and Expectations

Stimulus	Description	Loading
59	commander states...	-.76
55	enemy div. is alerted to...	-.73
36	mission is to...	-.73
52	aggressor expects that...	-.71
61	...is preparing to...	-.70
54	16CAA authorizes...	-.61
41	...at this time	-.60
42	...immediately	-.55
43	...in the future	-.54
35	...is capable of...	-.50

Two factors seemed to be concerned with sources of information. Evidently, at least in the opinion of the analysts involved, information source constitutes meaningful evaluative information to the intelligence analyst. One of these factors, factor X, seemed concerned with sources extraneous to the field situation which are possibly unreliable (e.g., "civilian indicates," "radio broadcasts indicate"). The second of these information source factors, factor XVI, was concerned with information derived from written materials. Factor X was named Extra-Military Information Sources, while factor XVI was named Contents of Documents. The stimulus items loading heavily on each of these factors are presented respectively in Tables 2-29 and 2-30.

Table 2-29

Stimulus Items Loading Heaviest on Factor X--
Extra-Military Information Sources

Stimulus	Description	Loading
50	**civilian** indicates...	.82
47	**radio** broadcasts indicate...	.72

Table 2-30

Stimulus Items Loading Heaviest on Factor XVI--
Contents of Documents

Stimulus	Description	Loading
46	**aggr.** document indicates...	-.75
16	**document** originates from 16CAA	-.64

Factor XI loaded well on only one stimulus item--"increases." This **factor** is essentially uninterpretable and may be an artifact of a degree of over **factorization.** Certainly, the stimulus item "increases" would logically seem **to** relate to factor II, Numbers, Trends, and Sources. In fact, the loading of **the** item "increases" on factor II was -.35. This loading tends to support the **interpretation** given. Alternatively, and as discussed above, the "trends and **numbers**" aspect could be removed from the definition of factor II and treated **as a** separate factor.

Deserter oriented information seemed to represent a unique considera**tion** to the subjects involved. This type of information was subsumed by factor **XII,** which was named Deserter Information. Information relative to supplies **(other** than armament) and equipment availability was embedded in factor XIII, **named** Logistics. The stimulus items loading heavily on these two factors are **presented** in Tables 2-31 and 2-32.

Table 2-31

Stimulus Items Loading Heaviest on Factor XII--Deserter Information

Stimulus	Description	Loading
49	deserter claims that...	.79
33	deserter entered friendly lines	.53

Table 2-32

Stimulus Items Loading Heaviest on Factor XIII--Logistics

Stimulus	Description	Loading
3	supplies are received	.45
60	...is logistical	.41

The two remaining factors, XIV and XV, were weak but probably meaningful. Factor XIV was named Unique Actions (e.g., CG enters maneuver area). This type of information may have factored separately because the entrance of the CG to an area possesses special implications. Such an action certainly is unusual and might warrant special consideration by the intelligence analyst. Factor XV, the final factor, seemed concerned with civilian-military relationships, as in the stimulus item "population cooperates with aggressor" (loading .49), and was named accordingly.

Discussion of Results of Factorization

The message sample has been manipulated into a set of factor groupings which are logically coherent and show some consistency across subjects. Apparently, expert judges were able to process the information carried by each message and converge on a similar structure for the message set. The resulting structure seems to possess construct validity for development of an organizational scheme for battlefield linguistically oriented information. Each subject in the sample of four persons permitted the production of possibly rival factor solutions. Repeated challenges to a variety of possible solutions were presented and systematically discounted.

As discussed relative to factor XI (page 37) and in view of the eigenvalues for the last extracted factor (page 30), there is some indication that slight over factorization may have been involved. In fact, some degree of over, rather than under, factorization was deliberately superimposed. It is easier to combine or eliminate factors than to develop new ones. If experience indicates the need to combine or eliminate factors, such combination/elimination can be readily accomplished. Moreover, it seems quite important that an information bank not lose information. An increase in the number of available categorization options would seem to decrease the potential for information loss.

We note that the application generality of the resultant factors is dependent on the generality of the message set employed as stimuli. While there is no easy way of knowing the generality of the message set employed, future tests of the scheme are planned. These tests will provide at least initial insight into the generality of the categorization scheme derived.

There has been at least one prior study which possessed goals which are similar to those involved in the present work (Miron, Patten, & Halpin, 1975). Because of methodological differences, it is difficult to compare the results of the Miron, Patten, and Halpin study with those of the present Applied Psychological Services' study. In the Miron, Patten, and Halpin study, the approach dictated that the emergent structure be entirely dependent on the adequacy of a dictionary used to process the messages involved. Accordingly, the factor groupings of the intelligence reports in the Miron, Patten, and Halpin study were based on overlapping concept categories occurring among the reports. The results of their factor analysis, consisting of five factors, were then employed in an attempt to confirm the major concept categories numbering some 17.

In the present study, the resultant factor groupings were less constricted. The factors grew out of and are a direct product of the messages. A categorization scheme was sought which is the result of the factor analysis, not the reverse. Whereas the Miron, Patten, and Halpin study sought confirmation of major category tags with the methodology of factor analysis, the logic of the current study required the prior factorial extraction as a basis for dictionary development. A comparison of the taxonomies emerging from the two studies is presented in Table 2-33. The present study seems to have isolated factors not captured by the Miron, Patten, and Halpin work. This could be the result of the different approaches used. Additionally, the present study tended to over factorization as a precaution against missing any vital categorization. Provided that effective cross referencing is employed, it seems that the more fragmented scheme of the present study would provide less data loss and more efficient information insertion and retrieval. Certainly, many of the factors of the present study which were not evidenced in the prior study seem to possess merit on the face of things, e. g., communication, armament preparations, and training, morale, and attitudes. More importantly, the present study includes all the factors of the Miron, Patten, and Halpin work plus other factors. Accordingly, it can be argued that the results of the present study have missed little of the prior study and add a degree of clarification.

The work described above suggests that the perceived structure of the military field messages analyzed can be represented by 15(16) orthogonal factors. The factors enlarge on and subsume the five factors isolated by the Miron, Patten, and Halpin work. Additionally, it seems that the present factors are sufficiently robust. Acceptable agreement was evidenced among the factorial structures yielded by the linguistic similarity perceptions of the four judges involved. The agreement seems to hold despite the wide experience diversity among the judges. It is possible that Army intelligence training tends to homogenize linguistic perception. While the resultant factorial structure seems generalizable to friendly activity, the actual veridicality of such generalization is not known.

Table 2-33

Comparison of Miron, Patten, and Halpin Factors with Factor of Present Study

Present Study		*Miron et al. Study*	
Factor	*Description*	*Factor*	*Description*
I	Movement and capability of movement	I,II	Large troop movements of strategic importance; Vehicular and small scale troop movement
II	Number, trends, and source	I,II	Large troop movements of strategic importance; Vehicular and small scale troop movement
III	Type of activity		———
IV	Training, morale, and attitudes		———
V	Communications		———
VI	Armament preparation		———
VII	Small arms capability	III	Unusual small arms fire
VIII	Strength and location of units	I	Large troop movements of strategic importance
IX	Plans and expectations		———
X	Extra military information sources	V	Civilian and POW reports of low reliability
XI	Eliminated		———
XII	Deserter information	IV	Deserter and POW reports of low reliability
XIII	Logistics		———
XIV	Unique actions		———
XV	Civilian military relationships		———
XVI	Contents of documents		———

Factorial Congruency: A Question of Logic

The question of degree of factor matching among the four analysts of the present study may also be examined by establishing a standard for comparison. Rather than compare factorial structures of the individual analysts with each other two at a time, each individual structure might be compared with the mean of the group. Since each individual involved in such a comparison is also a member of the group, such comparisons are statistically biased toward achieving a good fit. Nonetheless, such an analysis was considered useful at least as a check on the obtained results relative to the commonality of structure across the four analysts who provided the data reported and analyzed in the present chapter. To achieve the standard structure, the raw similarity matrices of the four analysts were averaged, correlational similarities computed (Stone et al., 1970), and the resultant single matrix was submitted to factor analysis and rotation. Sixteen factors were extracted. Each previously derived individual factor solution was compared with the average of the group, making a total of four comparisons of the individual subjects with the group mean. Accordingly, five data sets were involved in the four comparisons. Each comparison was accomplished in accordance with the cosine rotation technique developed by Kaiser, Hunka, and Bianchini (1971) and the agreement index developed by Butzkamin and Pfeiffer (1973). Since each factor analysis resulted in 16 factors, a 16 x 16 matrix of 256 cosines (representing the relationship of all rotated factors of one analyst with the group average) resulted for each of the four comparisons. As in the previous analyses, a number of controversies involving the identification of congruent factors resulted. The number of comparisons involved in each case ranged from 17 to 19. Here, the number of comparisons exceeding 16 roughly indexes the degree of controversy relative to goodness of fit achieved. Tables 2-34 through 2-37 present the resultant indices of agreement. The values of Tables 2-34 through 2-37 tend to be higher, on the average, than the previous individual comparisons.

To obtain the average agreement of all factors of each analysis with all factors based on the average of the group, the indices of agreement (Ai) were averaged across the total number of potentially corresponding factors. Table 2-38 presents the results of this averaging. The average indices of agreement presented in Table 2-38 ranged from 0.76 to 0.85, and are somewhat higher than the average of the previous comparisons among individuals. These data suggest strong agreement of the individual factorial structures with the group mean and therefore support the previous finding of strong agreement among the factorial structures of data yielded by the four separate subjects when compared with each other two at a time. A summary of the matched factors is presented in Table 2-39. The table of results suggests that with some small degree

Table 2-34 .

Agreement (A_i) Among Potentially Congruent Factors Produced by the Average of Group Versus Analyst M^c Alone

AVERAGE	M^c	A_i
1	7	0.9272
2	1	0.8828
3	2	0.9872
4	10	0.9653
5	4	0.8735
6	6	0.9433
7	9	0.9858
8	3	0.9463
9	13	0.9576
10	5	0.9366
11	15	0.9501
12	14	0.3960
13	14	0.5998
14	8	0.8830
15	12	0.6692
16	16	0.8811
15	11	0.6568
	MEAN	0.8495

Table 2-35

Agreement (A_i) Among Potentially Congruent Factors Produced by the Average of Group Versus Analyst M Alone

AVERAGE	M	A_i
1	2	0.9659
2	14	0.7143
3	4	0.9438
4	10	0.9662
5	11	0.9795
6	3	0.9788
7	9	0.8442
8	13	0.9302
9	1	0.9335
10	6	0.8875
11	15	0.8562
12	7	0.8372
13	5	0.8397
14	7	0.8284
15	8	0.7714
16	14	0.6136
7	12	0.5837
16	16	0.6110
	MEAN	0.8380

Table 2-36

Agreement (Aj) Among Potentially Congruent Factors Produced by the Average of Group Versus Analyst K Alone

AVERAGE	*K*	*Ai*
1	1	0.8469
2	4	0.9060
3	2	0.9561
4	3	0.8474
5	11	0.9158
6	6	0.8732
7	10	0.9226
8	5	0.6582
9	9	0.6947
10	4	0.5946
11	12	0.7805
12	13	0.8101
13	5	0.6199
14	16	0.7411
15	14	0.7829
16	9	0.8193
16	7	0.5535
4	8	0.6283
12	15	0.7516
	MEAN	0.7738

Table 2-37

Agreement (Aj) Among Potentially Congruent Factors Produced by the Average of Group Versus Analyst S Alone

AVERAGE	*S*	*Ai*
1	15	0.9646
2	5	0.8265
3	4	0.8939
4	1	0.7183
5	2	0.6951
6	13	0.9649
7	6	0.9676
8	1	0.8652
9	3	0.7895
10	16	0.5987
11	11	0.6728
12	7	0.9178
13	9	0.7813
14	11	0.6043
15	12	0.8814
16	10	0.3923
10	8	0.5610
5	14	0.6574
	MEAN	0.7640

Table 2-38

Summary of Mean Agreements (A_i) Among Potentially Congruent Factorial Structures Produced by the Average of Aanalysts M^C M, K, and S Versus the Individuals Taken Separately

COMPARISON	*MEAN A_i*
Group Average vs. M^C	0.8495
Group Average vs. M	0.8380
Group Average vs. K	0.7738
Group Average vs. S	0.7640

Table 2-39

Summary of Matching Factorial Structures Based on the Averaged and Individual Perceptions of Analysts M^C, M, K, and S

	Factors by			
GROUP AVERAGE	*M^C*	*M*	*K*	*S*
1	7	2	1	15
2	1	14	4	5
3	2	4	2	4
4	10	10	3	(14)
5	4	11	11	2
6	6	3	6	13
7	9	9	10	6
8	3	13	5	1
9	13	1	(7)	3
10	5	6	(8)	16
11	15	15	12	11
12	(12)	7	13	7
13	14	5	(15)	9
14	8	(12)	16	(10)
15	(11)	8	14	12
16	16	(16)	9	(8)

NOTE: - Factors having questionable congruency with the group average have been circled.

of mismatching, the identification of corresponding factors across all analyses and factors can be achieved. Factors having questionable congruency with the group average are circled in Table 2-39. In all cases, the factors having questionable congruency do not exceed three out of the possible total of 16.

These findings support the earlier conclusion of strong agreement based on a complete pairwise comparison of individual perceptual structures with each other.

CHAPTER III

TAXONOMY DEVELOPMENT

Chapter I of the present report described the logic, methods, and results of the steps completed to derive the factors underlying battlefield language. These factors, it is held, can constitute a partial basis for a taxonomy which can structure and categorize battlefield messages.

A taxonomy is defined as "a classification of data according to their natural relationships or the principle governing such classification" (English & English, 1959) or as "a way of simplifying a complicated universe of individual events and objects according to some useful way of identifying the way in which groups of individual (observations) have things in common or differ" (Miller, 1967).

Taxonomic schemes have been found useful in other contexts. The physical sciences, life sciences, and behavioral sciences are three areas in which systems of organizing data, objects, or ideas into groups or categories have been found useful. It is possible that military intelligence information, with its basis in the language system, can profit from a taxonomy which would lead to increased information understanding and interpretive ease.

When considering such taxonomies, it is important to distinguish between taxonomies of attributes and taxonomies of individuals (Green & Carmone, 1970). The taxonomies used to describe individual differences in the behavioral sciences are usually classifications of attributes of people, and less often classifications of the people themselves, such as sex and race. In biological taxonomies, on the other hand, the elements are categories composed of the biological organisms themselves, such as maple, pine, and oak. While such typological classifications of species in biology have survived, similar attempts by behavioral and social scientists to develop classification systems have often lacked permanency. This may be an artifact of the generalization that social and behavioral phenomena more often fall between the idealized types than fit them.

Regardless of the possible opaqueness of a taxonomy for battlefield messages, we note that such taxonomies are held to be useful by others. In this regard, Sells (1963) wrote in another context:

> The most obvious need is for the development of a taxonomy and measurement technology of variables describing the stimulus situation....But no comparable dimensions of the stimulus situation have been systematically studied, and consequently only piecemeal and usually inadequate account can be taken of variance attributable to the situation (p.5).

Miller (1967) separated rigorous and nonrigorous taxonomies. Miller contended that all taxonomies are nonrigorous except those which are dependent on an actual count or measurement. Miller set up the following objectives for any taxonomy which will be useful:

- limited to 15 to 20 classes
- classes can be learned in a few hours
- characterizations should foster anticipation of types of errors
- level of detail should be no greater than possible prediction accuracy

Under Miller's conception, the present taxonomy would fall in the "nonrigorous" category and the present taxonomy would meet the first two (and possibly the third) of Miller's objectives. The fourth objective is not highly relevant in the present case since the purposes of the taxonomy are descriptive rather than predictive.

The Taxonomy

In order to form the desired taxonomy, the 15 interpretable factors derived in Chapter I and summarized in Table 3-1 were considered to constitute major battlefield message descriptive attributes. However, since the factors grew from the linguistically derived kernel and operator classes, it seemed best not to neglect the content analytic classes in which the factors were rooted. Accordingly, the second axis of the taxonomy was considered to be the classes from which the factors were derived. Because 48 classes seemed to represent excessive detail and, in some case, undesired partial redundancy, the 48 classes were reduced to 27. These are listed in Table 3-2.

Table 3-1

List of Factors in Taxonomy

1. Movement and capability for movement
2. Number, trends, and source
3. Type of activity
4. Training, morale, and attitudes
5. Communication
6. Armament preparation
7. Small arms capability
8. Strength and location of units
9. Plans and expectations
10. Extra military information sources
11. Deserter information
12. Logistics
13. Unique actions
14. Contents of documents
15. Civilian-military relationships

Table 3-2

List of Classes in Taxonomy

1. Composition of units
2. Markings and insignia
3. Position of forces and equipment
4. Strength of units
5. Replacements
6. Quantitative indications
7. Weapons and equipment
8. Analysis
9. Characterization of missions
10. Sabotage
11. Actual and threatened contact
12. Strategy and tactical doctrines
13. Preparations and readiness
14. Movement
15. Chronology
16. Information about officers
17. Competence of personnel and equipment
18. Mental and physical condition of troops
19. Civilian-military relations
20. Sources/methods of obtaining information
21. Informer source-military
22. Informer source-nonmilitary
23. High reliability source
24. Low reliability source
25. Conditional phrases
26. Unknown information
27. Other

The taxonomy may be conceived as a matrix with the 15 factors forming one axis and the 27 classes forming the second axis. Row-column intersects yield cells which are called categories. Any message may be assigned to one or more categories. For example, the message "Infantry are trained in individual combat" would fall in the category defined by the intersect of the "Training, Morale, and Attitudes" factor and the "Competence of Personnel" class. The advantages of such a taxonomy are that it could:

- provided a basis for a battlefield message categorization scheme
- help to assure complete categorization of any message
- help to assure complete recapture of any message
- provide a basis for profile analysis of message content to yield an index of change over time
- provide a structured basis for further research into battlefield language and for certain aspects of analyst training
- provide the basis for quantitative analysis of an intelligence data base and of analyst decision making
- form the substrate for at least a part of a battlefield information management scheme
- help the battlefield observer prepare clearer, improved, more comprehensible messages
- highlight the relationships among various types of message content
- improve communications accuracy
- improve message content since the messages would be cast against the backdrop of a common set of factors
- help to standardize battlefield language and language and language interpretation
- provide a master filing system of battlefield language

The principal immediate application of such a taxonomy, assuming required additional development work, is a computer based support system which would facilitate the processing of intelligence data. A prior study (Miron, Patten, & Halpin, 1975) suggested the use of computers in such contexts to relieve the user of clerical activities and to provide analytic and interpretive tools for transforming information into intelligence.

For such a system, both the taxonomy and the data organization are crucial. Both will affect the utility and meaningfulness of the system to users. The taxonomy can provide a set of constructs. However, unless the constructs are presented in a structured and understandable form, they may possess little practical utility.

Dictionary Development

Having derived a taxonomy, the next step appeared to be the development of a scheme which would provide some basis for providing reliable and accurate use of the taxonomy. Such a scheme seems required in view of the richness of language and the wide range of words which are used almost interchangeably. To this end, a "dictionary" (Appendix B) was developed. The dictionary was conceived as a multipart compendium which would allow quick, accurate, and reliable use of the taxonomic system.

The dictionary contains five major sections--factor definitions, class definitions, compendium of messages, uniterms, and isonyms.

The initial two sections, factor definitions and class definitions, contain definitions of the factors and the classes included in the taxonomy.

The third section contains a compendium of messages which are relevant to each factor/class (category). The messages included are the 110 which formed the basis for the factor analysis. Each category is illustrated by a set of messages. The compendium of messages may be helpful in several ways: (1) the messages included in any category help to define that category, (2) a user with a new message could match the new message to one of the included messages; this would help to assure proper taxonomic classifications.

The fourth section, the index of uniterms, consists of a list of alphabetically arranged key terms, each followed by factor and class indicators. The user is thereby referred to the specific factors and classes that contain information relative to a specific key term.

If a taxonomy user wished to process very specific information such as "equipment," "sabotage," or "documents," reference to these terms in the uniterm section would direct him to the appropriate factor(s)/class(es) containing the information. The specificity of the uniterm index allows a savings in time, since only specific factor(s)/class(es) may be examined rather than an entire factor or an entire class.

The final dictionary section, the index of isonyms (iso = same; nym = word), is much like a thesaurus providing semantic equivalents for various related battlefield words. The isonym list also references each listed word to the taxonomic category to which the word refers. The user is accordingly supplied with synonomous key terms, as they appear in the message compendium section of the dictionary.

The dictionary, in its current form, is experimental, exploratory, and exemplary. Essentially, the dictionary may be viewed as a training or job aid--offering information on how to use the taxonomy and where to find information classified in accordance with the taxonomy.

The data base on which the dictionary, in its current form, was developed was small. If the convenience and value of the concept can be demonstrated, it can be expanded to include a greater range of messages and associated materials. It may be that a taxonomic classification of the messages relating to a given battlefield situation slights the continuous instability of such situations. A battle never is; it is always becoming. In addition to a wide variety of other information, a total message management scheme would need to recognize the constant change of flux while accepting the utility of categorical analysis.

CHAPTER IV

VALIDATION

Because the factors of the taxonomy were derived empirically and because the classes were derived from a structured linguistic analysis, the taxonomy seems logically defensible. However, the internal constructs of such a taxonomy tell one little about the pragmatic value of the taxonomy. One test of pragmatic value, in the present case, is whether or not intelligence analysts can reliably use the taxonomy. To develop initial insights into this question, two experiments were performed. In the first experiment, Army personnel with experience in intelligence were asked to read a set of battlefield scenarios. They classified messages relevant to the scenarios using a portion of the derived dictionary and taxonomy. In the second experiment, persons with similar experience were asked to read a set of scenarios and to state the taxonomic classifications in which they would expect to find information called for by questions relevant to the various scenarios.

Experiment 1--Information Classification

The purpose of experiment 1 was to test the ability of experienced intelligence persons to classify messages in accordance with the derived taxonomy. Since the total taxonomy involves more factors and classes than could be managed in the time available for data collection, a sample of factors and classes was selected for inclusion. The factors and classes sampled were:

Factors

A. Movement and Capability for Movement
B. Training, Morale, and Attitudes
E. Communications
J. Extra Military Information Sources
K. Deserter Information
L. Logistics

Classes

3. Position of Forces and Equipment
4. Strength of Units
6. Quantitative Indications
7. Weapons and Equipment
10. Sabotage
11. Actual and Threatened Contact
16. Information about Officers
18. Mental and Physical Condition of Troops
23. High Reliability Source
24. Low Reliability Source

These were selected to represent a varied type of content and a wide range of potential message categorizations. The abbreviated taxonomy contained a possibility of 60 categorizations (6 factors x 10 classes = 60).

Methods

Data were collected in a group session at the U. S. Army Intelligence Center and School, Fort Huachuca, Arizona. After assembling the group, a brief introduction in the purposes of the data collection was given. Then a formal training session in the use of the dictionary/taxonomy was presented. Briefly, the instruction consisted of an explanation of the total taxonomy, an explanation of the use of the dictionary, and practice in message classification. A large briefing-type display of the taxonomy was posted in front of the room during the training and used as an aid. The display remained posted during the formal data collection. This training consumed 45 minutes, and at the end of this time the analysts seemed comfortable with the dictionary/taxonomy.

After the training session, stimulus booklets were distributed, and the analysts were asked to work through the stimulus items. The analysts were allowed to use the dictionary in any manner they wished to assist them in their task. Completion time for all stimulus items and an associated opinion questionnaire was, on the average, 2 hours and 15 minutes.

Stimuli

The stimuli consisted of simulated battlefield messages and associated descriptive scenarios. The messages were either taken verbatim from Army training scenarios or modeled after these sources. Two sample messages are given in Table 4-1.

Table 4-1

Sample Stimulus Items

Message	Targeted Category*	Alternative Relative Categories*	Complexity Level
1. Deserter states his unit was rumored to be on a suicide mission	K-24	-	low
2. Civilians indicate that 14 tank units in the 1 F div maneuver areas as of 12 Sep are equipped with fording equipment. There are also amphibious tanks included in the force.	A-7	A-3, A-6 J-3, J-6, J-7	high

*Factor and class.

The set of 60 messages was designed to meet two specifications. Each message was targeted at one of the 60 factor-class combinations (called categories) included. The information contained in the first sample message in Table 3-1 can be described by the targeted factor-class pair, K-24 (where K stands for the factor "deserter information" and 24 stands for the class "low reliability information source"). Randomly, one-fourth (15) of the messages were chosen for use with each of four types of scenarios. The messages and their associated scenarios were structured to agree with one another in terms of time, day, date, location, and general type of activity. A second variable, complexity level of the message, was also varied. Complexity refers to the number of categories to which a message might pertain: 1 or 2 for low complexity; 3, 4, or 5 for medium complexity; and 6 or more for high complexity. The first sample message in Table 4-1 was designed to be relevant to only one of the 60 categories. The second sample message was designed to be relevant to six categories (including the targeted category A-7) as indicated in Table 4-1.

Scenario Development

The scenario associated with each message provided a context within which the analysts could evaluate the associated battlefield information. Each scenario was about 200 words in length and consisted of a description of a current situation, what led up to it, and the types of problems which are currently on hand. One scenario was developed to exemplify each of the following types of battlefield situations: attack, retreat, holding action, and logistics. Since there were 60 message stimuli and four scenarios, each scenario was referenced to 15 messages. Five of the 15 messages for each scenario were randomly chosen and modified to be of low complexity, another 5 were randomly chosen and modified to be of medium complexity, and the remaining 5 were modified to be of high complexity.

Response Forms

Two methods of recording the categorization(s) of the simulated battlefield messages were employed: (1) a multiple choice approach with all 60 messages, and (2) a free choice approach with a subset of these messages. The distractors for the multiple choice aspects were randomly chosen from those categories not appropriate to a given message. The free choice method allowed the analysts to choose a limited number of categories which they considered to be appropriate to a given message. The free choice response form presented the factors and classes in a matrix form. The analysts were

asked to indicate their response(s) by marking the appropriate cells. The number of choices that the analysts were asked to make using this matrix form was equal to the number of categories potentially relevant to the particular message. Accordingly, the analysts were advised of the possible number of category choices for each message categorized using the matrix answer form. The matrix response form appeared for each item employing this format as follows:

Factors

Classes	A	B	E	J	K	L
3						
4						
6						
7						
10						
11						
16						
18						
23						
24						

The task of the analyst was to state, within a limited number of choices, which factor-class intersect(s) (category) the presented message most logically fell.

The free choice matrix form was originally planned to be the sole method of recording categorizations. However, pilot work indicated this method to be quite time consuming. Therefore, the multiple choice approach described above was developed. The matrix form allowed an assessment, at least in part, of the complete categorization choices of the analysts. Also, due to time limitations, analysts were not asked to rate the categories chosen from their first choice (or most relevant) to their last choice (least relevant). An example of the multiple choice response form is presented as Figure 4-1. which shows a portion of the instruction page included in each stimulus booklet. The parallel instruction page for the matrix response is presented as Figure 4-2.

EXAMPLE: Assume that you have read a scenario about a battlefield situation in which the friendly forces are engaged in a holding action. Consider the following message and category choices:

Message	Category Choices
SPOTTED APPROX 40-50 AGG TROOPS PROCEEDING THROUGH WOODS VIC PA592798 AT 151227.	E10 O D3 O A3 O A16 O K6 O

Confidence Score _____

Of the categories given to the right of the message, category A3 (movement and capability for movement/position of forces and equipment) most closely typifies the message. (Verify for yourself that the other categories are not appropriate.) Therefore, you would darken the circle next to A3. The choice of A3 here seems to be clearly correct. Accordingly, you would select a confidence score of "7" and enter your response ("7") in the space provided.

The message with the responses filled would therefore look like this:

SPOTTED APPROX 40-50 AGG TROOPS PROCEEDING THROUGH WOODS VIC PA592798 AT 151227.	E10 O D3 O A3 ● A16 O K6 O

Confidence Score 7

Messages 1-7 refer to scenario I. Read scenario I now. Then categorize messages 1-7 following the instructions above.

Figure 4-1. Partial example of multiple choice item type response instructions.

Section 2

Section 2 also contains sets of messages and each set is again related to a specific scenario. To complete the questions of each set first read the scenario. Then for each message:

a) Select one category which the message seems to fit. Write a "1" in the cell corresponding to that category.

b) Select a second category which the message seems to fit. Write a "2" in the cell corresponding to that category.

c) Continue making selections and write the number of each (3,4,5 and so on) in the appropriate cells. Limit your number of choices to the limit indicated above the answer grid.

d) Use the dictionary to help you, if you want to.

Example: To further clarify this procedure we will classify a sample message. Assume that you have read a scenario about a battlefield situation in which the friendly forces are engaged in a holding action. Consider the following message and response matrix.

Message

Spotted approx 40-50 agg troops proceeding through woods VIC PA592798 at 151227.

Limit your choice of categories to 3.

Factors

Classes	A	B	E	J	K	L
3						
4						
6						
7						
10						
11						
16						
18						
23						
24						

Figure 4-2. Partial example of matrix response item type instructions.

For each scenario type, one low, one medium, and one high complexity question was randomly selected to be answered using the matrix method.

An additional dependent variable, confidence in categorization choice, was measured for various messages categorized. After each multiple choice categorization, the analysts recorded their confidence in the accuracy of their categorization using a seven point scale. The categorization scheme employed was:

1 = Less than 5% confident that selected category is correct
2 = Between 5% and 20% confident that selected category is correct
3 = Between 20% and 40% confident that selected category is correct
4 = Between 40% and 60% confident that selected category is correct
5 = Between 60% and 80% confident that selected category is correct
6 = Between 80% and 95% confident that selected category is correct
7 = More than 95% confident that selected category is correct

Test Booklets A and B

Each analyst received only a portion of the 60 messages to categorize. A random sample of 30 of the multiple choice message categorization items were assigned to one booklet (A). The remaining 30 multiple choice items were assigned to a second booklet (B). All booklets contained the 12 matrix form message categorization items.

Opinion Questionnaire

In addition to the 30 multiple choice and 12 matrix response type items, each booklet contained a final section which sought to obtain the opinions of the subjects relative to the taxonomy. The opinion questionnaire contained 12 items:

1. an open ended question relative to the parts of the system which were easiest and hardest to use

2. a three choice categorical question relative to the amount of detail included in the factors

3. an open ended question asking for examples of where the factors are too detailed or underdetailed

4. a magnitude estimate of the "naturalness" of the factors

5. a magnitude estimate of the "understandability" of the factors

6. a magnitude estimate of the "comprehensiveness" of the factors

7. an open ended question asking for any factors which should be added or deleted

8. same as question 2, but relative to classes

9. same as question 3, but relative to classes

10. same as question 4, but relative to classes

11. same as question 5, but relative to classes

12. same as question 6, but relative to classes

13. same as question 7, but relative to classes

14. a final open ended question asking for "additional comments"

Stimulus Development

To develop the necessary scenarios and associated messages, a former intelligence officer first developed the required scenarios. Then, the messages were fit to the scenarios. The former intelligence officer then reviewed the materials for naturalness and reality. Next, the materials were tried on a small pretest sample. The results of this pretest yielded areas of required revision and indicated the need to shorten the materials to meet administrative time constraints. This was accomplished, and the final package was again reviewed by the former intelligence officer. Finally, the booklets were produced in their final form.

Participants

The participants were 29 Army officers in the initial stages of the military intelligence staff officers advanced course at the U. S. Army Intelligence Center and School, Fort Huachuca, Arizona. Approximately 90 per cent were Captains and the remaining 10 per cent were First Lieutenants. The mean time in the Army of these men was 77.6 months, with an average of 59.7 months experience in intelligence. Sixteen analysts completed stimulus booklet A and 13 completed booklet B.

For analysis, the data of the two groups were combined to reconstitute the original pool of 60 multiple choice items.

All analysts received those items in which messages were to be classified into the 6 x 10 matrix representing the intersection of factors/classes.

Results

The single most important measure was considered to be the amount of agreement among analysts in selecting the alternatives. This was measured by kappa, an index of agreement developed by Fleiss (1971). Kappa is defined as the proportion of agreeing pairs of raters out of all possible pairs of raters. This technique provides both individual message agreement (P) as well as overall agreement (k), which is simply the mean of the individual P values.

According to Fleiss, the extent of agreement among the n raters for the ith item may be indexed by the proportion of agreeing pairs out of all the n(n-1) possible pairs. This proportion is:

$$P_i = \frac{1}{n(n-1)} \sum_{j=1}^{k} n_{ij}(n_{ij}-1)$$

$$= \frac{1}{n(n-1)} \left(\sum_{j=1}^{k} n_{ij}^2 - n \right)$$

where: n = number of raters
n_{ij} = number of raters who assigned the ith item to the jth category
k = number of categories of the nominal scale

The overall extent of agreement may then be measured by the mean of the P_i values:

$$\bar{\bar{P}} = \frac{1}{N}\sum_{i=1}^{N} P_i$$

$$= \frac{1}{Nn(n-1)}\left(\sum_{i=1}^{N}\sum_{j=1}^{k} n_{ij}^2 - Nn\right)$$

where: N = total number of items

Some degree of agreement is expected solely on the basis of chance. The correction factor, as developed by Fleiss (1971), such that the overall agreement corrected for the amount of agreement expected by chance is kappa (k):

$$k = \frac{\bar{\bar{P}} - \bar{\bar{P}}_e}{1 - \bar{\bar{P}}_e}$$

where: $\bar{\bar{P}}_e$ = mean proportion of agreement if the assignments were made at random

Fleiss (1971) also provided formulas for calculating the standard error of kappa. This allows test of the hypothesis that k = 0.

To illustrate the meaning of $\bar{P}$ (proportion of agreement on an individual item) and kappa (the mean of the separate agreements) an example is presented in Table 4-2.

Note that the actual $\bar{P}$ values and kappa tend to present a conservative point of view. For item 2, in which five of six judges are in agreement P is only .66 and for item 3 on which four of six judges agreed on one option and two judges agreed on another option, the P value is below .50. It is evident that P and k are more complex than a simple proportion. They are proportions involving permutations in the numerator and denominator and the total number of possibilities (the denominator) is sometimes misleadingly large. For example, for each item there are n(n-1) possible pairs of agreements (6 raters taken 2 at a time = $\frac{n!}{(n-r)!}$ = n(n-1) = 6(5) = 30 possibilities. For item 1 in Table 2, there are n(n-1) or 6(5) = 30 agreements out of a possible 30 agreements and P = 1.00. For item 2 there are n(n-1) or (5)(4) = 20 agreements out of a possible 30 and P = .66. For item 3 there are n(n-1) + n(n-1) or (4)(3) + (2)(1) = 14 agreements out of a total 30 and P = .466, etc.

Table 4-2

Example of P and k Values Among Six Raters on Assigning Ten Items to One of Five Alternatives

Item	Alternative 1	2	3	4	5	$\sum_{1}^{k} n_{ij}^2$	$\sum_{1}^{k} n_{ij}^2 - n$	$P_i = \frac{\sum n^2 - n}{n(n-1)}$
1	6	-	-	-	-	36	30	1.000
2	5	1	-	-	-	26	20	.666
3	4	2	-	-	-	20	14	.466
4	4	1	1	-	-	18	12	.400
5	3	3	-	-	-	18	12	.400
6	3	2	1	-	-	14	8	.266
7	3	1	1	1	-	12	6	.200
8	2	2	2	-	-	12	6	.200
9	2	2	1	1	-	10	4	.133
10	2	1	1	1	1	8	2	.066

$$\overline{P} = \frac{1}{N}\sum_{1}^{N} P_i = \frac{3.797}{10} = .3797$$

$$k = \frac{\overline{P} - \overline{P}_e}{1 - \overline{P}_e} = \frac{.38 - .20}{1 - .20} = \frac{.18}{.80} = .225$$

We note that the P for item 3 is less than .50 in spite of what seems to be considerable agreement. However, note that only 14 pairs of agreement occurred out of the possible 30. The P values decrease rather rapidly as the amount of agreement decreases until at item 10, where the worst agreement is illustrated. This item might represent the actual possibility for pure random effects of assignment by the six judges. Consequently, this P value would be appropriate for use as an adjustment for random effects on agreement for each item.

The obtained P value for each multiple choice item, by scenario, is presented in Tables 4-3 through 4-6, along with the corresponding raw agreement data. Failure to respond or missing data were treated as lack of agreement. Accordingly, the most stringent measures of agreement was used. Agreement on each of the items was statistically significant with all but three of the 60 items being significant at or below the .01 level. Agreement for each of the scenarios was also statistically significant at the .01 level. The overall uncorrected kappa value was .738 and the overall corrected kappa value was .664. These are quite high and represent considerable practical significance as well as statistical significance.

Table 4-3

Number of Agreements and P Values Among Judges on Selection of Alternatives for Each Message of Scenario 1

	Alternative					
Message	*1*	*2*	*3*	*4*	*5*	*Pi*
1	—	16	—	—	—	1.000
2	1	—	15	—	—	.875
3	—	—	—	15	1	.875
4	1	14	1	—	—	.758
5	—	—	16	—	—	1.000
6	—	15	—	—	1	.875
7	—	2	—	—	14	.767
8	6	1	—	1	5	.321*
9	—	13	—	—	—	1.000
10	—	6	5	1	1	.321*
11	—	11	—	—	2	.718
12	13	—	—	—	—	1.000
13	—	—	—	13	—	1.000
14	—	—	—	12	—	.846
15	11	1	1	—	—	.705

k= .804 *k corrected for chance= .749*

*Statistically significant at the .05 level. Values above .321 are statistically significant at or below the .01 level.

Table 4-4

Number of Agreements and P Values Among Judges on Selection of Alternatives for Each Message of Scenario 2

	Alternative					
Message	*1*	*2*	*3*	*4*	*5*	*Pi*
1	1	11	—	—	4	.510
2	—	16	—	—	—	1.000
3	—	10	—	4	2	.430
4	—	12	—	—	4	.600
5	3	—	12	1	—	.575
6	—	16	—	—	—	1.000
7	—	16	—	—	—	1.000
8	—	—	—	16	—	1.000
9	4	5	—	4	—	.282*
10	1	12	—	—	—	.846
11	—	3	—	8	2	.410
12	3	1	—	9	—	.500
13	—	—	2	1	10	.590
14	—	2	—	7	3	.314*
15	2	10	—	—	1	.590

k= .643 *k corrected for chance= .542*

*Statistically significant at .05 level. Values above .314 are statistically significant at or below the .01 level.

Table 4-5

Number of Agreements and P Values Among Judges on Selection of Alternatives for Each Message of Scenario 3

	Alternative					
Message	1	2	3	4	5	Pi
1	1	=	15	=	=	.875
2	11	=	4	1	=	.510
3	=	=	=	16	=	.750
4	=	1	15	=	=	.875
5	4	6	=	=	6	.300
6	2	14	=	=	=	.760
7	1	4	=	=	11	.508
8	1	=	=	12	=	.846
9	1	12	=	=	=	.846
10	1	3	=	=	9	.500
11	=	13	=	=	=	1.000
12	13	=	=	=	=	1.000
13	=	2	2	7	=	.282*
14	4	7	2	=	=	.359
15	=	=	1	=	12	.846

k= .684 *k corrected for chance= .595*

*Statistically significant at the .05 level. Values above .282 are statistically significant at or below the .01 level.

Table 4-6

Number of Agreements and P Values Among Judges on Selection of Alternatives for Each Message of Scenario 4

	Alternatives					
Messages	1	2	3	4	5	Pi
1	—	—	—	—	16	1.000
2	—	—	—	16	—	1.000
3	7	—	9	—	—	.475
4	—	—	16	—	—	1.000
5	15	—	—	—	1	.875
6	—	—	—	—	16	1.000
7	—	16	—	—	—	1.000
8	11	1	1	2	1	.467*
9	13	—	—	—	—	1.000
10	2	1	1	9	—	.474
11	13	—	—	—	—	1.000
12	12	—	—	1	—	.846
13	—	—	12	—	1	.846
14	—	3	9	—	—	.495
15	—	12	—	1	—	.846

k= .822 k corrected for chance= .771

*Statistically significant at the .01 level as are all values above .467.

An analysis of the variance among the P values was completed by scenario. The results indicated no statistically significant differences among scenarios. A summary of the results of this analysis is presented in Table 4-7.

Table 4-7

Summary of Between Scenario Variance Analysis (P Values)

Source	SS	df	MS	F
Between Scenarios	.3493	3	.1164	2.03
Within Scenarios	3.2077	56	.0573	
Total	3.5570	59		

Confidence

The amount of indicated confidence of the subjects in their choices is presented by item in Table 4-8. Translated into percentages, the overall mean confidence value was quite high. Because the confidence ratings were skewed towards the upper end of the scale, median confidence ratings were also calculated. The obtained median confidence level is presented below by scenario along with the corresponding mean value:

Scenario	Median	Confidence Level %	Mean	Confidence Level %
1	6.59	80-95	6.09	80-95
2	6.31	80-95	5.65	80-95
3	6.05	80-95	5.69	80-95
4	6.32	80-95	6.02	80-95
Overall	6.40	80-95	5.86	80-95

The obtained individual scenario median confidence levels reflect considerable confidence on the part of the subjects in their message categorizations.

Table 4-8

Mean (M) and Standard Deviation (S. D.) of Subjective Confidence by Message and Scenario

Message	Scenario 1		Scenario 2		Scenario 3		Scenario 4	
	M	SD	M	SD	M	SD	M	SD
1	6.60	.63	5.73	.22	6.13	1.13	6.30	.88
2	6.87	.36	6.47	.64	4.93	1.27	6.53	.64
3	6.40	.74	4.50	2.13	6.73	.46	6.31	.95
4	6.40	.99	6.40	.83	5.46	1.36	5.60	1.40
5	6.93	.26	5.73	1.03	5.53	1.36	6.33	.98
6	6.53	.74	6.73	.59	4.80	1.42	6.47	.52
7	6.60	.63	6.87	.35	5.27	1.53	5.93	1.28
8	4.53	1.56	6.93	.26	6.29	.99	6.00	1.25
9	6.93	.27	4.57	2.21	6.36	.74	5.21	2.01
10	4.29	1.68	6.80	.58	5.79	1.53	5.21	1.80
11	5.93	1.26	5.86	1.03	7.00	0.00	6.21	.80
12	5.71	1.98	5.29	1.38	6.00	1.80	6.07	.92
13	4.71	2.58	5.21	1.72	4.29	2.52	6.36	.93
14	6.57	.63	2.79	2.32	5.43	1.28	5.93	1.33
15	6.29	1.86	4.86	1.92	5.43	1.40	5.86	1.51
Overall Mean	6.09		5.65		5.69		6.02	
S.D.		1.28		1.35		1.37		1.22

To determine if agreement was related to confidence, the mean (Table 4-8) data were plotted and the correlation between agreement and confidence was calculated. The plot is presented as Figure 4-3. The resulting product moment correlation, .65, was statistically significant at the .01 level of confidence. The magnitude of the correlation coefficient suggests that mean degree of confidence varies directly with agreement on the selected alternative. The regression equation relating agreement to confidence is $Y' = 2.2343X + 4.2140$, where Y' is predicted item agreement and X is mean subjective confidence for the item.

An analysis of the variance of the mean confidence assignments among the four scenarios was completed. The results are summarized in Table 4-9, which indicates no statistically significant difference in confidence as a function of scenario type.

Table 4-9

Summary of Variance Analysis of Confidence Ratings

Source	SS	df	MS	F
Between Scenarios	2.2251	3	.7417	1.05
Within Scenarios	39.4988	56	.7053	
Total	41.7239	59		

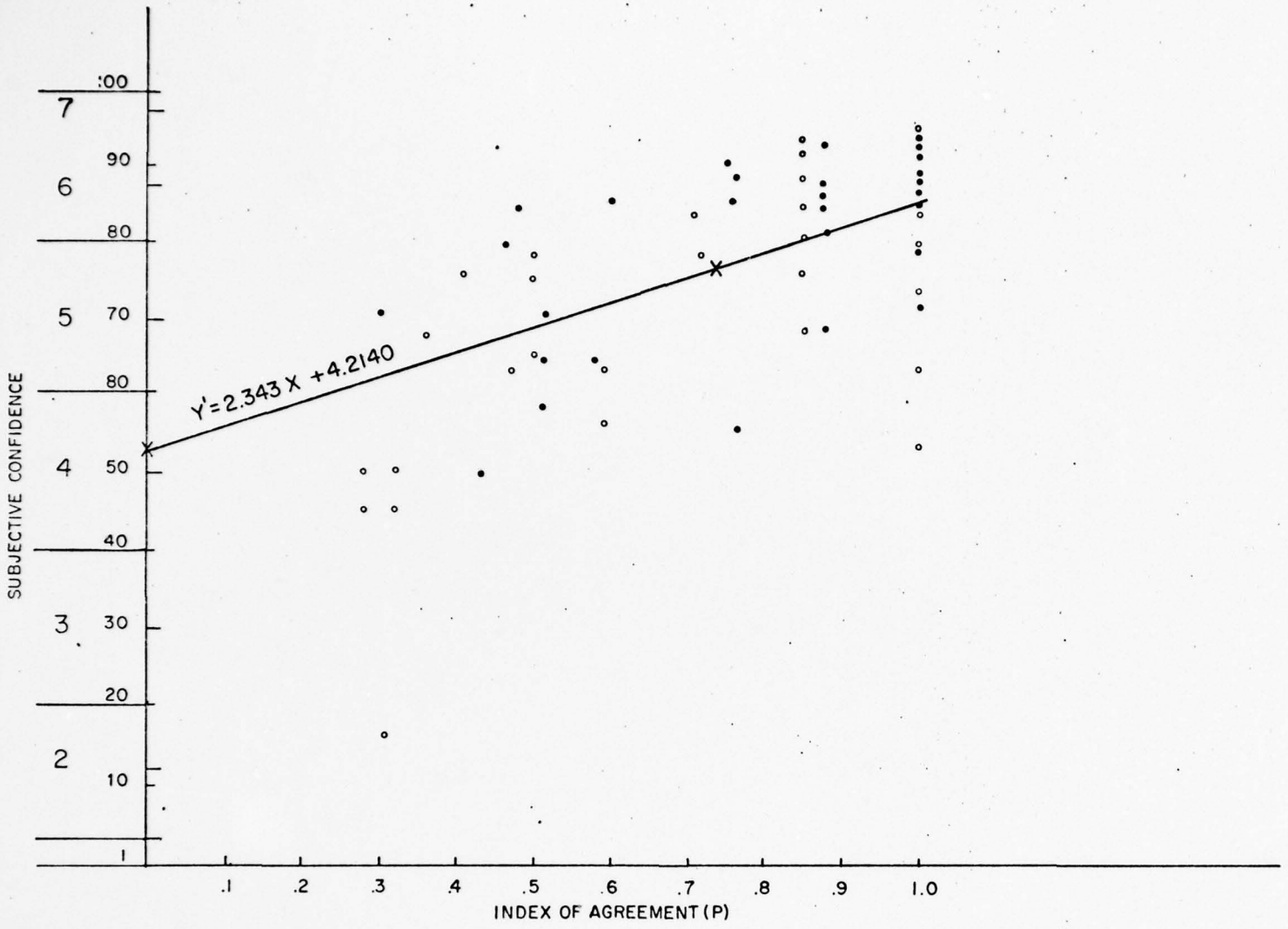


Figure 4-3. Subjective confidence as a function of overall agreement (dots = Form A; circles = Form B)

Matrix Classification

The matrix classification data were analyzed by extending Fleiss' (1971) formula from each judge making one assignment per item to each judge making more than one categorization per item. The derivation of this extension is presented in Appendix C. The resultant P values are presented in Table 4-10.

Table 4-10

P Value by Item for Matrix Classification

Item	Choices	P
1	3	.3034
2	2	.5787
3	6	.4626
4	1	.3943
5	5	.2359
6	10	.2448
7	2	.4764
8	3	.3529
9	6	.2897
10	1	.3333
11	3	.1888
12	6	.4021
Overall corrected for chance		.3552

All P values in Table 4-10 are statistically significant below the .01 level. There was a considerable loss of agreement when the alternatives were increased from 5 as in the multiple choice format to 60 as in the matrix (6 x 10) format. The extent of agreement, however, was still considerable when the characteristics of kappa are considered. The relationship between agreement and number of choices was determined. A nonstatistically significant correlation, .34, was indicated. These findings tend to imply that the multiple choice format (not the number of choices) tended to inflate the "agreement" level.

The matrices were further analyzed to determine whether factor or class was more important in determining placement of the item. Fleiss' basic technique was modified and used to determine the extent of agreement for the factor receiving the highest number of assignments and for the class receiving the highest number of assignments. The modification of kappa was simply the proportion of agreeing pairs for a factor (or class) with the most agreement for each item over the total possible number of pairs. This is a measure of agreement and can be interpreted similarly to kappa. The results are presented in Table 4-11 in terms of the factor and class receiving the highest number of agreements.

Table 4-11

Agreement by Factor and by Cl for Matrix Categorization

Item	Choice	Factor	Factor Agreement	Class	Class Agreement
1	3	A	.267	3	.132
2	2	A	.422	18	.561
3	6	K	.337	3	.218
4	1	E	.393	7	.395
5	5	E	.159	16	.073
6	10	A	.142	7	.059
7	2	K	.424	18	.455
8	3	J	.369	3	.246
9	6	J	.147	10	.199
10	1	J	.318	10	.315
11	3	E	.113	4	.089
12	6	J	.348	3	.161
		Mean	.287		.242

While Table 4-11 suggests a slight advantage for the factor agreement, the difference is not statistically significant when tested by the sign test (two tailed). Moreover, although the variability appears much greater for class agreement than for factor agreement, the variance ratio is not statistically significant ($F = 1.89$). This suggests that the hypothesis that factor selection stability is greater than class selection stability cannot be accepted.

Application of Theory of Signal Detection to Message Classification

In the first part of this experiment (multiple choice responses), the analysts were asked to classify messages into selected categories of the developed taxonomy. In this testing situation, the analyst's task became that of identifying the correct category from among the alternatives provided. One additional test of pragmatic value, in the present case, is whether or not intelligence analysts can effectively use and feel comfortable with the taxonomy. Alternatively, what is the analyst's operating characteristic? The evaluative construct used here comes from the theory of signal detection (Green & Swets, 1966; Swets, 1964). Since this approach is well known, it is only briefly considered in this report. In the signal detection literature, two types of experiment have been described (Egan & Clarke, 1966) to generate the receiver operating characteristic (ROC). One type of curve has been termed the stimulus conditional ROC and the other has been termed the response conditional ROC.

The stimulus conditional ROC is based on estimated probabilities P(Y/S) and P(Y/N) conditional upon the occurrence of signal (S) and noise (N). Alternatively, the yes (Y) or the no (N) response of the subject is conditional upon the events of signal and noise that give rise to the stimuli. According to Green and Swets (1966):

> Yes-no decisions are made by the receiver of a communication when he must take one or another action based on his degree of confidence in the correctness of his identification response. Similarly, when the receiver sends back his identification to the source, the source may be required to confirm or reject it. The analogy between correct configurations and false alarms, has led the analyses in terms of the operating characteristic. Although the analogy is not exact in the case of the receiver, based as it is on a discrimination between responses rather than between stimuli, the operating-characteristic analysis serves its major purpose (p. 313).

The response-conditional ROC is designed for situations in which the observer makes an identificatory response. This identification represents the observer's best judgment as to which event gave rise to the information received. To generate a response-conditional ROC, the subject must assign ratings of confidence to his identifications. The question answered by this procedure is how well the subject can partition his ratings of confidence between his correct and incorrect identifications. Accordingly, the response-conditional ROC represents the relation between the probability $P(Y/C)$ that the observer will accept his identification as being correct when it is in fact correct and the probability $P(Y/I)$ that he will accept his identification as correct when it is actually incorrect. Previous workers who have noticed the relevance of signal detection theory to performance on mental tasks include, but are not limited to, Miles (1973) and Siegel and Pfeiffer (1969).

The binary techniques typically associated with signal detection theoretic analyses were not employed here since more efficient use of data occurs when the rating method is employed. The relation between the binary-decision procedure and the rating method is that in the rating method used here, a response of "7" represents a "yes" under a strict criterion. If the observation is in fact correct, then $P(R_7/C)$ is equivalent to $P(Y/C)$ where R_7 is the assigned confidence rating "7." This is first computed. Similarly, for the incorrect intervals, $P(R_7/I) = P(Y/I)$. Next, the conditional probability of a response of "6" to a correct choice is computed. The following assumption is then made: observations accepted as correct under a given criterion will also be accepted under a less strict criterion. Therefore, the second value of $P(Y/C)$ is $P(R_7 + C) + P(R_6 + C)$. Thus, the value of $P(Y/C)$ corresponding to the criterion established by the rating is the cumulative probability:

$$P(Y/C) = \sum_{i=7}^{1} P(R_i/C)$$

The corresponding value of $P(Y/I)$ is found employing the same logic:

$$P(Y/I) = \sum_{i=7}^{1} P(R_i/I)$$

Computation of $P(Y/C)$ and $P(Y/I)$ continues over the entire rating scale ($i = 7, 6, \ldots, R$) and these data are employed to generate the receiver operating characteristic (ROC). A slight deviation from this described procedure was necessitated by the lack of data in the lower confidence

categories. Accordingly, the data for confidence categories 1, 2, 3 were pooled. This pooling reduced the number of total confidence categories from 7 to 5 as follows:

0.0 - .40	.40 - .60	.60 - .80	.80 - .95	.95 - 1.00
(1, 2, 3)	(4)	(5)	(6)	(7)

(Low) ← (Confidence Level) → (High)

The results of the calculations for the 16 analysts are presented in Table 4-12. The columns of Table 4-12 are headed by the confidence probabilities that define the five rating categories. The first section, consisting of the first row of the table, lists the frequency with which each rating category, R_i, was employed. The second section shows four quantities based on correct trials, and the third section of Table 4-12 shows the corresponding quantities based on incorrect trials.

Table 4-12

Results of ROC Analysis

CONFIDENCE LEVEL	*0.0-.40*	*.40-.60*	*.60-.80*	*.80-.95*	*.95-1.00*
(1) $f(R_i)$	16	33	63	118	250
(2) $f(C,R_i)$	6	24	45	103	238
$P(C/R_i)$	.375	.727	.714	.873	.952
$P(R_i/C)$	.014	.058	.108	.248	.572
P(Y/C)	1.000	.986	.928	.820	.572
(3) $f(I,R_i)$	10	9	18	15	12
$P(I/R_i)$	.625	.273	.286	.127	.048
$P(R_i/I)$	.156	.141	.281	.234	.188
P(Y/I)	1.000	.844	.703	.422	.188

NOTE: - (1) Section (2) is based on 416 observations
(2) Section (3) is based on 64 observations
(3) N= 480 observations

Before discussing the construction of the ROC curve, consider a calculation which displays more directly the intelligence analysts' ability to use the taxonomy. The first row in section 2 of Table 4-12, labeled $f(C, R_i)$, presents the frequency of the compounded occurrence of correct trials and the use of category R_i. The row immediately below it lists estimates of the conditional probability $P(C/R_i)$; specifically, it shows the proportion of the observations placed in each category that resulted from the occurrence of a correct trial. For example, the value 0.952 in the extreme right hand column is obtained by dividing the compound frequency (238) by the total number of observations placed in that category (250).

The quantity $P(C/R_i)$ increased directly with increases in the probabilities that define the confidence categories. This result is plotted in Figure 4-4 and indicates that the analysts employed the rating categories in a reasonable manner.

The last two rows in sections 2 and 3 of Table 4-12 illustrate the calculation of the proportion of response conditional hits and false alarms. The third row in each section, $P(R_i/C)$ and $P(R_i/I)$, is the reverse of the conditional probability previously considered. Thus, for example, the value 0.572 in the extreme right hand column of section 2 is obtained by dividing the number of correct identifications made in that category (238) by the total number of correct identifications (416). Finally, the fourth row in each section, the response conditional hit and false alarm proportions corresponding to each of the five decision criteria, is a cumulation of the quantities in the third row moving from right to left in the table. It represents the procedure of compressing rating data to "yes" - "no" by considering successively each of the four category boundries as defining a "yes" - "no" criterion.

According to Coombs, Dawes, and Tversky (1970), the performance of any sensory device, human or otherwise, operating on the basis of the theory of signal detectability (TSD) can be completely described by what is called the receiver operating characteristic (ROC) curve. Figure 4-5 presents the ROC curve for the 16 subjects of the present study. Each subject responded to 30 multiple choice items to yield a total of 480 observations. This value is near the minimum of 500 observations recommended by Green and Swets (1966) for such an analysis. The detection sensitivity of the analysts, $d' = 1.11$, is based on the mean d_i' of the highest four confidence categories (7, 6, 5, 4) presented in Table 4-12. Since the indices of discriminability ($d_7' = 1.06$, $d_6' = 1.12$, $d_5' = 0.94$, $d_4' = 1.33$) are nearly the same at all points, it seems that the ROC function for the subjects and materials employed is relatively bias free.

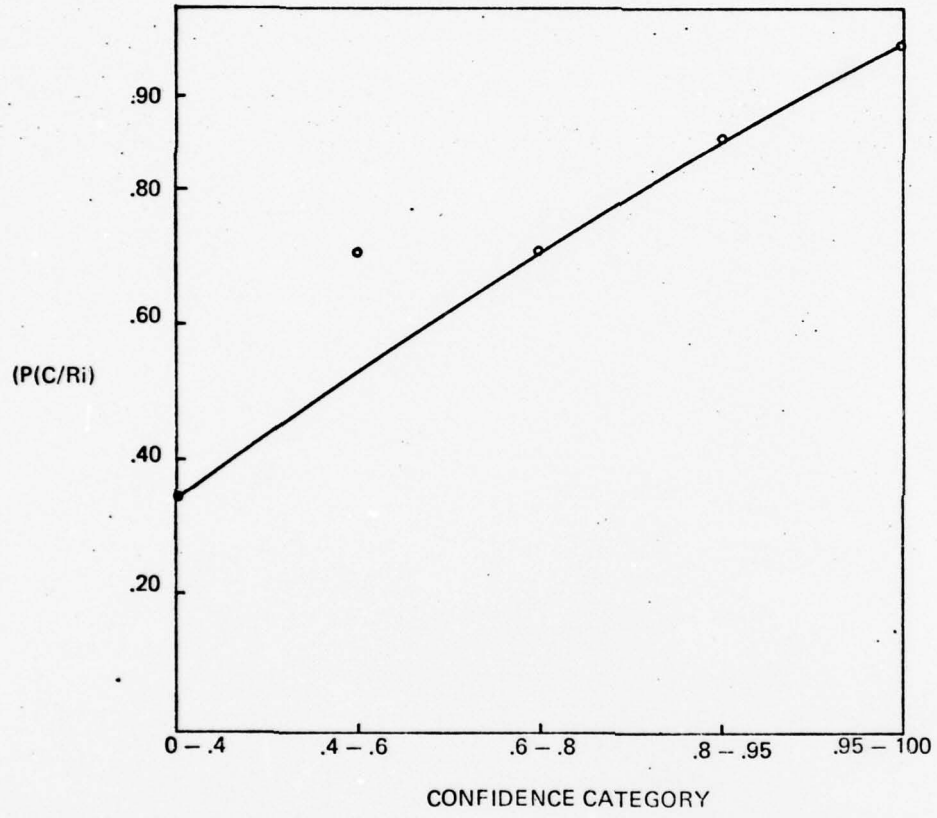


FIGURE 4-4. PROPORTION OF OBSERVATIONS PLACED IN EACH RATING CATEGORY THAT RESULTED FROM A CORRECT IDENTIFICATION

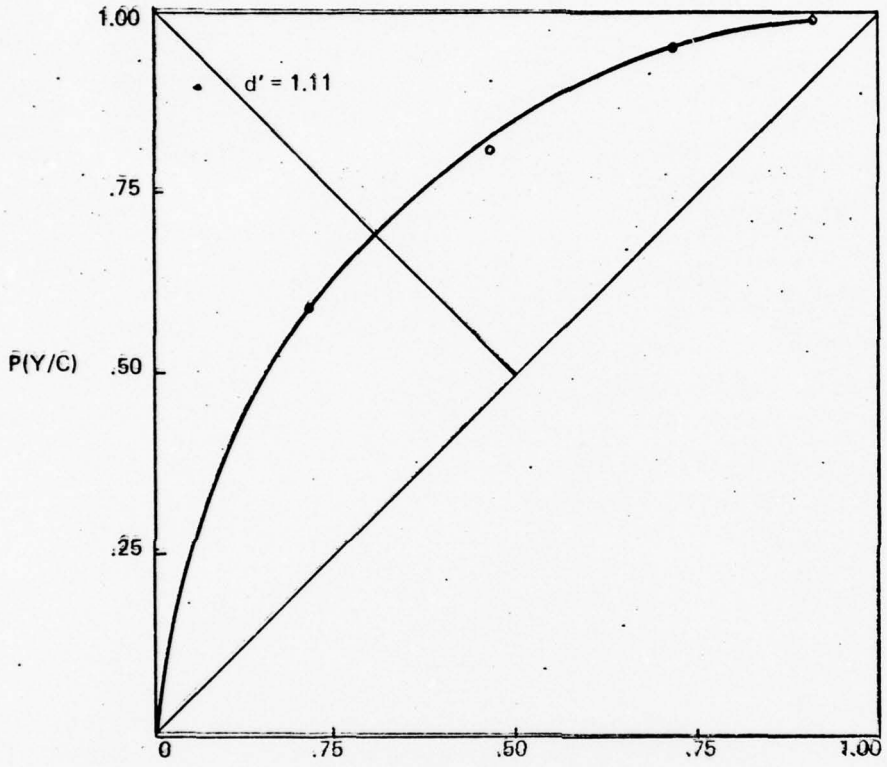


FIGURE 4-5. RESPONSE CONDITIONAL ROC CURVE FOR PRESENT SUBJECTS

The question answered by the ROC curve is how well the intelligence analyst can partition his ratings of confidence between his correct and incorrect identifications. Alternatively stated, the question is how sensitive the observer is to the quality of his own performance. It is apparent that the variability of the data points about the best fitting theoretical curve is very small.

Analyst confidence was directly related to correct message categorization and detection sensitivity was at a satisfactory level when using the taxonomy. Accordingly, the practical utility of the taxonomy is supported from the signal detection theoretic point of view, i. e., analysts are sensitive to the quality of their own performance.

Discussion

The results presented above support the contention that from the point of view of message classification, the derived taxonomy and associated dictionary possess considerable merit. High agreement was evidenced among users of the system for message classification purposes. This agreement seemed to sustain itself across scenario types and to hold for both the factors and the classes of the taxonomy. Moreover, even with limited training, the subjects were confident about their ratings, and were sensitive to the quality of their performance. This confidence also sustained itself across scenarios.

It is quite possible that the obtained agreement and user sensitivity would attenuate if the total taxonomy had been employed. However, the amount of this attenuation is a matter of conjecture at this junction.

Accordingly, from the point of view of message classification, the derived system seems to possess some merit.

Experiment 2--Category Query

Experiment 2 sought to investigate whether or not experienced battlefield intelligence persons could agree, with the assistance of the developed dictionary, on which taxonomic category would most likely contain the information to allow answering questions relative to battlefield situations.

As for experiment 1, only a portion of the total taxonomy was included. The factors and classes sampled in experiment 2 were:

Factors

C. Type of Activity
F. Armament Preparation
H. Strength and/or Location of Units
I. Plans and Expectations
N. Contents of Documents

Classes

1. Composition of Units
5. Replacements
8. Analysis
11. Actual and Threatened Contact
13. Preparations and Readiness
17. Competence of Personnel and Equipment
19. Civilian-Military Relations

These were selected to possess minimum redundancy with the factors and classes included in experiment 1 and to apply to a wide range of battlefield situations. In this case, the abbreviated taxonomy consisted of 35 categories (5 factors x 7 classes = 35).

Method

The methods paralleled those for experiment 1. Data were similarly collected at the U. S. Army Intelligence Center and School, Fort Huachuca, Arizona. The same training syllabus as in experiment 1 was followed with the exception that the practice was oriented towards information location and retrieval rather than towards message classification.

Stimuli

The stimuli consisted of 12 scenarios describing battlefield situations and questions containing battlefield conditions and activities relative to these scenarios. Two sample questions are given in Table 4-13.

Table 4-13

Sample Stimulus Items: Experiment 2

Question	Target Category
1. Are the aggressor forces in the area of PAO287 preparing for or capable of a major river crossing?	H-13
2. Where are the aggressor armor strengh concentrations in this sector?	F-8

Four sets of questions were assembled to be relevant to 5 factors and 7 classes. Each set of 35 questions contained one item targeted for each of the possible 35 factor-class pairs (categories). For example, the information needed to answer sample question number 1 in Table 4-13 is most likely to be found in targeted category H-13 (factor: Strength and/or Location of Units, class: Preparations and Readiness).

Response Form

The multiple choice approach was employed with all 140 simulated battlefield questions. The distractors for each multiple choice item were randomly chosen from those categories not appropriate to that question. A sample page of the response booklet is presented as Figure 4-6.

Questions 1-6 refer to scenario I. Read scenario I now. Then categorize questions 1-6 following the instructions.

Questions	Category Choices	
1. What units are engaged in fortifying the emplacements on hill 451 at the present time?	H11 O C17 O F1 O	I11 O C1 O
2. Are there any aggressor engineering units in the area or any plans or preparations for for fording the Weisse-Elster river?	H1 O I8 O I5 O	F1 O I1 O
3. Does the training of replacements confirm previous reports that the aggressor forces are capable of CBR warfare?	N1 O N17 O N11 O	I5 O I1 O
4. Have sightings revealed any changes in position of weapons with potential nuclear capacity?	H11 O I17 O H19 O	F17 O F11 O
5. Are there any indications that the aggressor aggressor forces will utilize their nuclear weapons against friendly units?	N11 O N17 O I1 O	I11 O I17 O
6. What are the general plans of the aggressor during the standdown period?	F1 O F8 O F5 O	H5 O I8 O

Questions 7-12 refer to scenario II. Read scenario II now. Then categorize questions 7-12 following the instructions.

7. Send G2 information on sightings of any weapons with nuclear capability?	I17 O H17 O H19 O	H13 O C17 O
8. Have the local civilians become aware of aggressor preparation for possible use of nuclear weapons?	C13 O F19 O F17 O	C19 O I5 O
9. Have any local farmers been recruited for guerrilla units?	I19 O F13 O H13 O	F5 O N8 O
10. Do documents found in hidden guerrilla supply cache (VIC PA432472) indicate the plans of the irregulars in this area?	H8 O C8 O N1 O	N13 O I1 O

Figure 4-6. Page of experiment 2 response booklet.

Test Booklets A and B

Each analyst received only a portion of the 140 questions to categorize. A random half of the questions for each scenario was assigned to one booklet (A). The remaining questions relevant to that scenario were assigned to a second booklet (B). The two forms of the test were organized into two versions of test booklets to facilitate the data collection. Fifteen analysts completed booklet A and 16 completed booklet B.

Feedback Questionnaire

The same opinion questionnaire as included in experiment 1 was included in experiment 2.

Scenario Development

Scenarios similar to those used in experiment 1 were developed to provide a context for the information processing task. Twelve scenarios, three scenarios of each of the four types used in experiment 1, were developed. Each scenario type was queried relative to all 35 targeted categories. The 35 targeted categories were rando[illegible] divided among the three scenarios of a type (with the constrain[illegible] the distribution of number of questions per scenario be: 12, 12, an[illegible]

Stimulus Development

Stimulus development paralleled that for experiment 1, including initial preparation by the same former intelligence officer, pretest, revision, and final assembly in the form of two booklets. Completion of each stimulus booklet, including the associated opinion questionnaire, took about one hour and 45 minutes.

Participants

The participants in experiment 2 were 30 captains and one first lieutenant enrolled in the military intelligence staff officers advanced course at Fort Huachuca. These officers had been in the Army for slightly more than 6 years, on the average. The range was 4 years to slightly over 12 years. On the average, the analysts had 4 years, 9 months experience in intelligence work. Two of the analysts had no intelligence experience, while the analyst with the greatest amount of military intelligence experience had 11 years and 8 months of such duty.

Results

The kappa statistic (Fleiss, 1971) was calculated to determine the extent of agreement among the analysts on the task of identifying that category which should be queried for obtaining information needed to answer a question relative to the scenario situation. The analyses were performed separately on Forms A and B. In addition, a kappa on the combined data (Form A plus Form B) was also obtained. The obtained values, corrected for chance kappa, were: Form A = .52; Form B = .43; overall = .48.

On the basis of chance alone, the following proportions of agreement would be expected: Form A = .21; Form B = .20; overall = .20.

Standard errors of kappa were obtained to examine the hypothesis of no agreement beyond chance. The obtained standard errors of kappa were: Form A = .006; Form B = .005; overall = .005. Calculation of the statistical significance of the obtained kappa values from these standard error values indicates that the obtained kappa values are significantly different from chance.

Kappa values were also calculated for each item by scenario. These statistics are presented in Table 4-14, which also presents the raw agreement data and the P values by item. Each of the 12 scenario kappa values presented in Table 4-14 is statistically significant at or below the .01 level.

Examination of the statistical significance of the agreement findings on individual items indicated that agreement on 17 items, those whose P value was less than .295, did not reach significance. The agreement on two items was significant at the .05 level; these were items whose P values were between .295 and .313. The agreement among the analysts on the remaining 121 items was statistically significant at or below the .01 level.

Table 4-14

Number of Agreements and P Values Among Judges on Selection of Alternatives for Each Item and Scenario

		Alternative					
	Item	*1*	*2*	*3*	*5*	*4*	*Pi**
Scenario I	1	—	—	14	—	1	.867
	2	—	15	—	—	—	.524
	3	10	1	1	—	3	.457
	4	—	—	1	2	12	.638
	5	5	1	—	9	—	.438
	6	3	—	10	1	2	.408
	7	3	4	—	—	9	.375
	8	—	5	—	9	2	.392
	9	6	2	—	3	5	.242
	10	—	—	2	12	2	.567
	11	1	—	—	—	15	.875

k= .395 *k corrected for chance*= .526

Scenario II	1	2	—	—	12	1	.638
	2	1	12	—	2	—	.638
	3	—	5	3	1	6	.267
	4	—	5	6	—	4	.295
	5	3	—	1	—	11	.552
	6	1	8	1	4	1	.324
	7	—	5	—	9	2	.392
	8	—	9	—	7	—	.475
	9	12	—	3	—	1	.575
	10	—	1	3	12	—	.575
	11	1	13	—	1	1	.650
	12	—	1	4	11	—	.508

k= .324 *k corrected for chance*= .491

Table 4-14 (cont.)

Scenario IV

Item	Alternative 1	2	3	4	5	Pi*
1	2	11	—	—	2	.543
2	—	1	4	4	6	.257
3	—	—	5	9	1	.438
4	—	1	—	14	—	.867
5	—	—	13	1	1	.743
6	15	—	—	—	—	1.000
7	6	1	3	2	4	.208
8	—	2	—	14	—	.767
9	1	—	11	1	3	.483
10	2	5	—	1	8	.325
11	1	6	—	—	9	.425

k= .432 *k corrected for chance= .551*

Scenario III

1	10	—	3	2	—	.467
2	14	—	—	—	1	.867
3	8	—	7	—	—	.467
4	3	—	12	—	—	.657
5	—	—	—	1	14	.867
6	—	15	—	—	—	1.000
7	—	—	2	—	13	.654
8	4	—	2	—	9	.354
9	6	3	6	—	—	.271
10	—	1	4	5	4	.175
11	8	1	—	—	6	.354
12	—	7	1	7	—	.346

k= 401 *k corrected for chance= .539*

Table 4-14 (cont.)

		Alternative					
	Item	1	2	3	4	5	Pi*
Scenario V	1	15	—	—	—	—	1.000
	2	1	4	6	1	3	.229
	3	—	1	13	1	—	.743
	4	—	14	1	—	—	.867
	5	—	3	—	11	1	.552
	6	6	1	5	1	2	.248
	7	7	1	—	—	8	.408
	8	2	1	5	8	—	.325
	9	—	8	—	—	8	.467
	10	—	1	11	4	—	.508
	11	5	1	8	2	—	.325
	12	5	9	—	2	—	.392

k= .371 *k corrected for chance= .505*

Scenario VI	1	—	1	10	—	4	.486
	2	—	—	—	15	—	1.000
	3	2	—	—	13	—	.752
	4	4	1	4	2	4	.181
	5	—	15	—	—	—	1.000
	6	—	—	15	—	—	1.000
	7	1	13	2	—	—	.658
	8	1	—	—	15	—	.875
	9	16	—	—	—	—	1.000
	10	—	—	16	—	—	1.000
	11	—	1	11	2	1	.463
	12	1	13	—	2	—	.658

k= .677 *k corrected for chance= .756*

Table 4-14 (cont.)

		Alternative					
	Item	1	2	3	4	5	Pi*
Scenario VII	1	—	1	—	1	13	.743
	2	4	11	—	—	—	.581
	3	1	—	1	13	—	.743
	4	6	4	3	2	—	.238
	5	2	10	2	—	1	.448
	6	16	—	—	—	—	1.000
	7	12	—	—	3	1	.575
	8	—	11	2	1	2	.475
	9	4	—	—	2	10	.433
	10	—	—	3	—	13	.675
	11	10	1	1	4	—	.425

k= .445 *k corrected for chance = .576*

Scenario VIII	1	1	1	12	—	1	.629
	2	4	—	5	—	6	.295
	3	1	—	—	12	2	.638
	4	—	14	—	1	—	.867
	5	1	—	10	—	4	.486
	6	6	1	4	—	4	.257
	7	—	1	—	8	7	.408
	8	11	2	1	1	1	.467
	9	1	6	—	6	3	.275
	10	15	—	1	—	—	.875
	11	—	1	15	—	—	.875
	12	5	1	3	—	7	.283

k= .402 *k corrected for chance= .530*

Table 4-14 (cont.)

		Alternative					
	Item	1	2	3	4	5	Pi*
Scenario IX	1	4	—	3	8	—	.352
	2	—	—	14	1	—	.867
	3	2	10	—	3	—	.467
	4	—	—	13	1	1	.743
	5	—	—	—	3	11	.548
	6	—	3	6	—	6	.314
	7	1	8	—	1	6	.358
	8	2	5	—	—	9	.392
	9	16	—	—	—	—	1.000
	10	14	1	—	—	1	.758
	11	1	1	3	6	5	.233
	12	—	1	1	—	14	.758

k= *k corrected for chance*= .566

Scenario X	1	2	—	12	—	1	.638
	2	—	—	1	11	3	.552
	3	—	1	—	1	13	.743
	4	—	—	15	—	—	1.000
	5	—	3	2	7	3	.267
	6	—	—	15	—	—	1.000
	7	—	—	14	1	1	.758
	8	1	—	3	3	8	.279
	9	—	—	2	—	14	.767
	10	2	3	1	10	—	.408
	11	2	11	—	3	—	.492

k= .489 *k corrected for chance*= .628

Table 4-14 (cont.)

Scenario XI

Item	Alternative 1	2	3	4	5	Pi*
1	—	—	—	—	15	1.000
2	3	—	12	—	—	.657
3	—	—	—	—	15	1.000
4	—	—	3	12	—	.657
5	—	7	8	—	—	.467
6	1	—	—	3	11	.552
7	1	13	—	1	1	.650
8	9	5	—	1	1	.383
9	—	1	—	5	10	.458
10	2	14	—	—	—	.767
11	—	1	—	13	1	.646
12	—	14	1	—	1	.758

k= .563 *k corrected for chance*= .666

Scenario XII

1	4	—	—	8	3	.352
2	3	1	11	—	—	.552
3	—	8	2	3	2	.314
4	1	10	2	1	—	.433
5	1	14	—	—	—	.867
6	—	11	3	1	—	.552
7	—	16	—	—	—	1.000
8	—	1	—	15	—	.875
9	—	—	—	8	8	.467
10	2	3	8	2	—	.271
11	1	1	—	2	12	.556
12	1	15	—	—	—	.875

k= .435 *k corrected for chance*= .593

*<.295; not statistically significant
>.295 <.313; statistically significant between .05 and .01 levels of confidence
>.313; statistically significant at or below .01 level of confidence

An analysis of the variance of the P values across scenarios was completed. A summary of the results of this variance analysis is presented in Table 4-15. The results indicate no statistically significant differences among the P values of the various scenarios.

Table 4-15

Summary of Between Scenario Variance Analysis (P Values)

Source	SS	df	MS	F
Between Scenarios	.7544	11	.0686	1.25
Within Scenarios	6.9727	127	.0549	
Total	7.7271	128		

Accordingly, the experiment 2 analysts demonstrated statistically significant agreement, and the extent of agreement did not vary across the 12 scenarios.

Agreement Between Experiments

The chi square technique was used to test the significance of difference between the kappa values of experiment 1 and experiment 2. Specifically, the question of concern was: does the agreement among the analysts in experiment 1 (whose task was to classify messages into the categorization scheme employed) differ significantly from the agreement found among the analysts in the second experiment (whose task was to identify categories they would select to obtain information required to answer intelligence related questions)? The overall kappas (across both forms in each experiment) were used in this analysis. The resulting chi square, 3.24, was not statistically significant. This indicates that although agreement tended to be higher among the analysts in the first experiment as compared with the analysts of the second experiment, the difference was not large enough to indicate that the two task types are independent in the population considered.

Response Correctness

The previous analyses pertained to the agreement among the Army intelligence personnel sampled on the identification of a category they would be likely to query if they were searching for specific intelligence data. To determine insight into response accuracy, the responses to the combined 70 questions were scored on an absolute scale for correctness of response. The scoring key was based on the responses of the research team, staff members of Applied Psychological Services. The mean number of correct responses (across both forms) was 49 or 70 per cent, with a standard deviation of 6.46. The range of correct responses was from 31 to 61 (44 to 87 per cent).

Kendall's coefficient of concordance was applied to the accuracy scores of the analysts of experiment 2 as they occurred on each of the 12 scenarios. The coefficient, calculated in this manner, provides an initial index of systematic variation in accuracy as a function of scenario. The resultant coefficient of concordance (corrected for tied ranks) was .106. This is interpreted as indicating minimum correlation, if any, between scenarios and accuracy.

To verify the prior interpretation, an analysis of variance was performed on the accuracy data to determine if a statistically significant difference existed among the accuracy scores for the scenarios. The results of the analysis are summarized in Table 4-16.

Table 4-16

Summary of Analysis of Variance for Scenarios (Accuracy Scores)

Source	SS	df	MS	F
Between Scenarios	12,869.26	11	1,169.93	2.83*
Within Scenarios	148,867.29	360	413.52	
Total	161,736.54	371		

* Statistically significant at the .01 level.

The F value, 2.83, was statistically significant at the .01 level. Tukey's gap test was applied to the data to determine those scenarios which were significantly different (Table 4-17).

Table 4-17

Gaps Between Scenario Means

		MEAN								
SCENARIO	MEAN	62	63	66	68	69	70	71	73	74
6	86	(24)	(23)	(20)	(18)	(17)	(16)	(15)	(13)	(12)
11	74	(12)	11	8	6	5	4	3	1	
3	73	11	10	7	5	4	3	2		
9	71	9	8	5	3	2	1			
8	70	8	7	4	2	1				
4	69	7	6	3	1					
12	69	7	6	3	1					
1	68	6	5	2						
7	68	6	5	2						
10	66	4	3							
5	63	2								
2	62									

Following Tukey's procedure, a D of 11.55 (significant at the .05 level) was obtained. A comparison of the differences of the scenario means with D indicated that the scenario 6 scores were significantly different from the other scenarios. Scenario 6 possessed the highest mean accuracy score. A second gap was found to exist between scenario 11 and scenario 2. No other statistically significant gaps were isolated. In view of the large number of possible gaps when 12 mean values are involved, the finding of only a limited number of gaps is interpreted as indicating that scenario type did not materially affect accuracy. Examination of the individual scenarios indicated no obvious differences between scenario 6 and the other scenarios or between scenarios 11 and 2.

These accuracy results are promising from the point of view of the ease of handling the categorization system employed in the study. Still better and more uniform performance might be anticipated with additional experience. In view of the minimal training involved, the analysts were able to perform with moderate accuracy, were again self consistent among themselves, and querying accuracy did not seem to be seriously debilitated by scenario type.

Discussion

The results of experiment 2 tend to be in conformity with those of experiment 1. Again, substantial agreement was evidenced among analysts and the agreement was scenario independent. Additionally, adequate accuracy was demonstrated in experiment 2, and the accuracy was largely scenario independent. When the experiment 2 results are coupled with the results of experiment 1, it seems that the derived system can be considered to be self contained both from the points of view of message classification and category inquiry. The coupled results support contentions favoring the possible potential of the derived system.

We note that although the difference between the overall kappa values for the two experiments was not statistically significant, the trend was towards higher agreement in experiment 1 than in experiment 2. The reason(s) for this trend are not entirely obvious. The inquiry task seems more opaque than the classification task, and this might suggest some need for additional system clarification relative to inquiry. It is also possible that qualitative differences existed in the training for the two groups. While the same training syllabus was followed for both groups, different administrators were involved. This was necessitated because the scheduling of the participants required that both data acquisition sessions take place at the same time. While the two administrators "rehearsed" together and attempted to "standardize" as much as possible, the conjecture of across administrator differences cannot be rejected. Additionally, the possibility of group differences remains open. However, this possibility does not seem tenable. The subjects were assigned by intelligence school personnel to the two groups in such a manner which seems to assure random assignment.

Other explanations for the findings relative to the trend implying differences between the inquiry and the retrieval tasks remain open. The "tests" for each type of task were not equated for difficulty, different scenarios were involved, and a larger number of scenarios was involved in experiment 2. For example, experiment 1 was based on four scenarios representing each of the following types of activity: retreat, attack, logistics, and holding. Three of each of these types of scenario were involved in experiment 2. Additionally, the categorization scheme was drawn directly from messages--not from questions. Accordingly, one might anticipate greater ease for direct categorization of messages than for questions relative to messages. Finally, it is possible that the two types of tasks rest on different mental operations or on different levels of the same mental operation. It seems that the data bank query task of experiment 2 demands more data evaluation than does the input task of experiment 1. Moreover, an analyst may possibly read more implications into a question than into a concrete message, and he may perceive more potential relationships.

From the point of view of the experimental results, it seems that a viable system has at least been partially achieved. However, as with any evolving system, additional clarification may be required for maximum utility. We note that the methods of study and analysis, here employed, were largely unique. There were no models to rely on and study, and analysis methods were developed in parallel with the taxonomic system. The methods developed here can form at least a partial foundation for future work of the sort which was the concern of the two experiments.

Opinion Questionnaire Results

After the completion of the data acquisition forms for each experiment, the analysts completed an opinion questionnaire consisting of 14 items on various aspects of the classification system. Some of the items were open ended, some were multiple choice, and others called for magnitude estimation.

Factors

When asked whether the factors: (a) were overly detailed, (b) had the proper amount of detail, and (c) were not detailed enough, the majority (59 per cent) of the analysts in experiment 1 (whose task involved the categorization of intelligence messages) indicated that the factors had the "proper amount of detail." A fair proportion of the sample, 33 per cent, thought the factors should be "more detailed." Only two analysts (8 per cent) indicated that they considered the factors to be "overly detailed." The analysts of experiment 2, on the other hand, were almost equally divided over the three options. These results suggest that for the task of categorizing messages, the factors are sufficiently detailed, but that additional factors might be helpful for the inquiry task. We note that the quantitative kappa values indicated less agreement among the analysts who performed the experiment 2 inquiry task as compared with those who performed the experiment 1 message classification task. This higher state of disagreement may be mirrored in the opinions of the experiment 2 analysts relative to the amount of detail included in the present system.

Very few of the analysts of either experiment took the opportunity to criticize constructively the factors when given the opportunity to do so in item 3. Some of the responses obtained from those who were critical of the factor detail were: too complex, overly detailed, armament preparation is not distinguishable from other factors, inadequate factors, and factor E (communications) should be separated into friendly and enemy communications.

Items 4, 5, and 6 involved rating the factors along scales which ranged from zero to 100, on naturalness, understandability, and comprehensiveness. The resultant means and standard deviations are presented by experimental group and characteristic in Table 4-18.

Table 4-18

Mean (M) and Standard Deviation (S. D.) of Ratings of Factors on Naturalness, Understandability, and Comprehensiveness

Characteristic	Experiment 1		Experiment 2	
	M	S. D.	M	S. D.
Naturalness	59. 8	18. 3	42. 2	20. 1
Understandability	66. 2	19. 3	50. 9	23. 2
Comprehensiveness	66. 2	19. 3	51. 2	27. 5

The respective frequency distributions of responses are presented in Tables 4-19, 4-20, and 4-21.

Table 4-19

Frequency Distribution of the Ratings of Naturalness of Factors

	Experiment 1	Experiment 2
0- 20	-	3
21- 40	8	15
41- 60	5	6
61- 80	14	5
81-100	1	-

Table 4-20

Frequency Distribution of the Ratings of Understandability of Factors

	Experiment 1	Experiment 2
0- 20	-	2
21- 40	4	10
41- 60	4	8
61- 80	17	8
80-100	3	1

BEST AVAILABLE COPY

Table 4-21

Frequency Distribution of the Ratings of Comprehensiveness of Factors

	Experiment 1	Experiment 2
0- 20	1	6
21- 40	3	9
41- 60	9	2
61- 80	11	8
81-100	3	4
No Response	2	2

Generally, the understandability and comprehensiveness of the factors were rated higher than the naturalness. This finding is difficult to comprehend because the factors are derived from a multidimensional scaling analysis of the message perceptions of intelligence officers. Nevertheless, on the average, the naturalness was rated at about the midpoint of the scale. For the experiment 1 analysts, the understandability and comprehensiveness were rated well above the scales' midpoints on the average, while for the experiment 2 analysts, these two characteristics were rated at about the midpoint. All three Table 4-19, 4-20, and 4-21 frequency distributions indicate the ratings of the experiment 1 analysts to be skewed towards the upper end. No experiment 1 analysts found the factors to be very unnatural or very difficult to understand. The skewness is not evident in the corresponding data for the experiment 2 analysts. Again, the increased difficulty evidently experienced by the experiment 2 analysts during their inquiry task seems to have generalized and affected their ratings of the factors.

About one-third of the sample of each experiment responded to item 7, which asked for suggestions relative to other factors that could be added to the taxonomy and for suggestions relative to factors which might be deleted without hindering the completeness of the factor set. The suggestions for additional factors included:

- allied forces information from front line troops
- terrain considerations
- weather reports
- special weapons
- air activity
- nuclear activity
- enemy intelligence collection
- personalities
- weapon systems

Two of the factors were identified as candidates for deletion. One analyst thought that the Plans and Expectations factor might be eliminated on the grounds that it is very close in meaning to the class called Preparations and Readiness. The other factors nominated for deletion by each of two subjects were Armament Preparation and Type of Activity. No reasons were given for these suggested deletions.

Classes

A parallel set of questions was included relative to the taxonomic classes. An exceedingly large proportion of the analysts of experiment 1 regarded the classes as either having the proper amount of detail or as being overly detailed (85 per cent). Three-fourths of the analysts in experiment 2 were of the same opinion. It appears that for the vast majority of the subjects, the classes, as they stand, incorporate adequate detail.

In a followup question (item 9), one analyst indicated that classes 23 and 24, High Reliability Source and Low Reliability Source, were difficult to understand. Two analysts suggested a consolidation of classes 13 (Preparations and Readiness) and 17 (Competence of Personnel and Equipment), which appeared to them to have overlapping meaning. The final constructive criticism was related to class 11, Actual and Threatened Contact. The analyst indicated that additional examples could help to explain the class more clearly. He gave the following examples: fire fighting, sighting, hearing, and indirect fire. On balance, it appears that minor class revisions along the lines suggested above would make them more acceptable to the military personnel involved in the present sample.

The naturalness, understandability, and comprehensiveness of the classes were evaluated using the same type of magnitude estimation scales as for the factors. The resultant means and standard deviations are presented in Table 4-22. The respective frequency distributions of responses are presented in Tables 4-23, 4-24, and 4-25.

BEST AVAILABLE COPY

Table 4-22

Mean (M) and Standard Deviation (S. D.) of Ratings of Classes on Naturalness, Understandability, and Comprehensiveness

Characteristic	Experiment 1		Experiment 2	
	M	S. D.	M	S. D.
Naturalness	64. 4	17. 4	46. 8	22. 2
Understandability	66. 6	17. 1	51. 0	24. 9
Comprehensiveness	64. 5	16. 5	56. 1	27. 3

Table 4-23

Frequency Distribution of Ratings of Naturalness of Classes

	Experiment 1	Experiment 2
0- 20	-	3
21- 40	5	12
41- 60	10	6
61- 80	11	7
81-100	2	1
No Response	1	2

Table 4-24

Frequency Distribution of Ratings of Understandability of Classes

	Experiment 1	Experiment 2
0- 20	-	3
21- 40	4	9
41- 60	8	8
61- 80	12	7
81-100	4	2
No Response	1	2

Table 4-25

Frequency Distribution of Ratings of Comprehensiveness of Classes

	Experiment 1	Experiment 2
0- 20	-	4
21- 40	3	4
41- 60	10	7
61- 80	12	9
81-100	2	4
No Response	2	3

Generally, the three characteristics considered were rated moderately above the midpoint of the scale by the experiment 1 analysts and at about the midpoint by the experiment 2 analysts. The respective frequency distributions reflect the same trend. No experiment 1 analyst rated the classes as very unnatural, very difficult to understand, or fully uncomprehensive. Yet, three or four of the experiment 2 analysts considered the classes to possess these extreme, negative characteristics. About half of the analysts of both groups rated the classes above 60 on all three characteristics. All of this seems to suggest that the taxonomic classes possess a moderate degree of acceptability but that further refinement would possess value from the point of view of fuller reputability. Yet, there is nothing in the current data which suggests that the class axis of the developed taxonomy has not at least partially achieved its objectives.

Evaluation of the System

Items 1 and 14 of the opinion questionnaire were open ended in nature and asked about the easiest and most difficult aspects of the system and for general comments. There were mixed opinions relative to the easiest and most difficult aspects of the system. While some indicated that they found it easy to classify the messages into classes and difficult to classify them into factors, about an equal number found the opposite to be the case. Very few analysts said that the system was difficult to use. The nature of the difficulties found with the system, as identified by the analysts in response to item 1, consisted of the following: classifying a message into a single category (factor/class), understanding the definitions, learning a new vocabulary, and being able to distinguish between factors and classes.

Additional comments received in response to item 14 did not indicate an unworkable or unacceptable system. A small number of analysts indicated that they would have liked more detailed definitions. Some analysts were of the opinion that certain factors and classes were not completely distinguishable from others and that adding more detail to the definitions might correct this situation. One analyst claimed that the system was too extensive for daily use and would be easier to use if it was made simpler.

In summary, while differences existed among the overall means of both groups on the ratings of naturalness, understandability, and comprehensiveness of the derived system, the ratings were generally above or at the midpoint of a 100 point scale. This suggests that a generally acceptable system is on hand. Additional and/or improved training, dictionary enhancement, and refinement of the taxonomic system might serve to augment the acceptability of the system.

CHAPTER V

MODEL INTEGRATION AND FUNCTIONAL DESCRIPTION OF COMPUTER IMPLEMENTATION OF THE BATTLEFIELD MESSAGE CLASSIFICATION SYSTEM

The prior chapters of the present report described the methods and results of work oriented towards the development of a taxonomy which describes the battlefield language structure. While the derived taxonomy was suggested to be adequately general and useable by battlefield analysts for message classification and retrieval purposes, the opinion questionnaire (Chapter IV) indicated concern by some analysts that the model was overarticulated in some aspects and underarticulated in other aspects. Accordingly, a total model review session was conducted at the U. S. Army Intelligence Center and School, Fort Huachuca. The purpose of the review was to acquire the insights of experienced intelligence personnel relative to areas of needed taxonomy extension and consolidation. The review was conducted informally and was performed by 14 officers (including one foreign military liaison officer) and senior enlisted personnel. All were currently active in some phase of battlefield information analysis.

The comments of these panel members were in general conformity with the indications of the Chapter IV opinion questionnaire results. On the basis of these comments, the taxonomy was revised and set in a revised current form. This revised taxonomy is called the "final model" in subsequent aspects of this report.

The Final Model

The final model is also conceived as a matrix with factors forming one axis and classes forming the second axis.

The factors of the final model are the same as in Table 3-1 of this report. However, the number of classes is reduced considerably. The classes removed and the associated reasoning are presented in Table 5-1.

PRECEDING PAGE BLANK-NOT FILMED

Table 5-1

Clases Removed from Original Taxonomy and Associated Reasoning

CLASS	NAME	REASON FOR REMOVAL
6	Quantitative Indications	Part of Factor 2--Number, Trends, and Source
11	Actual or Threatened Contact	Part of Factor 3--Type of Activity
14	Movement	Part of Factor 1--Movement and Capability for Movement
18	Mental and Physical Condition of Troops	Part of Factor 4--Training, Morale, and Attitudes
19	Civilian-Military Relations	Same as Factor 15--Civilian-Military Relationships
20	Source/Methods of Obtaining Information	Part of Factor 10--Extra Military Information Sources
21	Informer Source-Military	Part of Factor 10--Extra Military Information Sources
22	Informer Source-Civilian	Part of Factor 10--Extra Military Information Sources
24	Low Reliability Source	Part of Factor 10--Extra Military Information Sources

Additionally, four classes were added:

Terrain
Weather
Special Weapons
Air Activity

Accordingly, the classes in the final model are those presented in Table 5-2.

Table 5-2

Classes in Final Model

Composition of units	Chronology
Markings and insigna	Information about officers
Position of forces and equipment	Competence of personnel and equipment
Strength of units	High reliability source
Replacements	Conditional phrases
Weapons and equipment	Unknown information
Analysis	Terrain
Characterization of missions	Weather
Sabotage	Special Weapons
Strategy and tactical doctrines	Air Activity
Preparations and readiness	Other

Functional Description of Computer Implementation of the Battlefield Message Classification System

As pointed out by Finley, Muckler, Gainer, and Obermayer (1975), the development of a taxonomy which describes the natural order of things is not a sufficient end, in itself, to the practitioner. In the present case, the taxonomy was developed in order to make evident those battlefield language structure properties which should be employed as a basis for increasing battlefield information flow and intelligence emphasis. To this end, a functional design was derived of a system for processing battlefield messages using a general purpose digital system and the derived taxonomy.

We conceive of a message processing system composed of a network of typewriter/display terminals, manned by combat (military) personnel who represent both the sources of input messages for storage as well as by intelligence analysts who refer to and extract correlations from such messages. The network of terminals is deployed in a battlefield (or test) situation. The computer may be interfaced to the network through a variety of means. These are not of particular importance here. Figure 5-1 presents such a potential system for illustrative purposes.

The functional capabilities of the system given in the next section include storage of messages and message headers by class(es) and factor(s), a flexible user syntax, including inquiries, security features, automatic purging for archival storage, and user displays for teaching the operational procedures.

A block diagram of the program is presented in Figure 5-2, which shows the overall logic flow between the program modules. The composition of files to be stored is presented in Table 5-3.

Of particular importance in defining such a terminal oriented system is specification of the types of input messages which a user will have at his disposal. A brief list of the types of transactions considered is given in Table 5-4. These capabilities are further detailed in Appendix D in the form of user syntax diagrams for each transaction type (input, maintenance of the store of messages and the networks, inquiries, status display requests, and user orientation (teaching) capabilities.

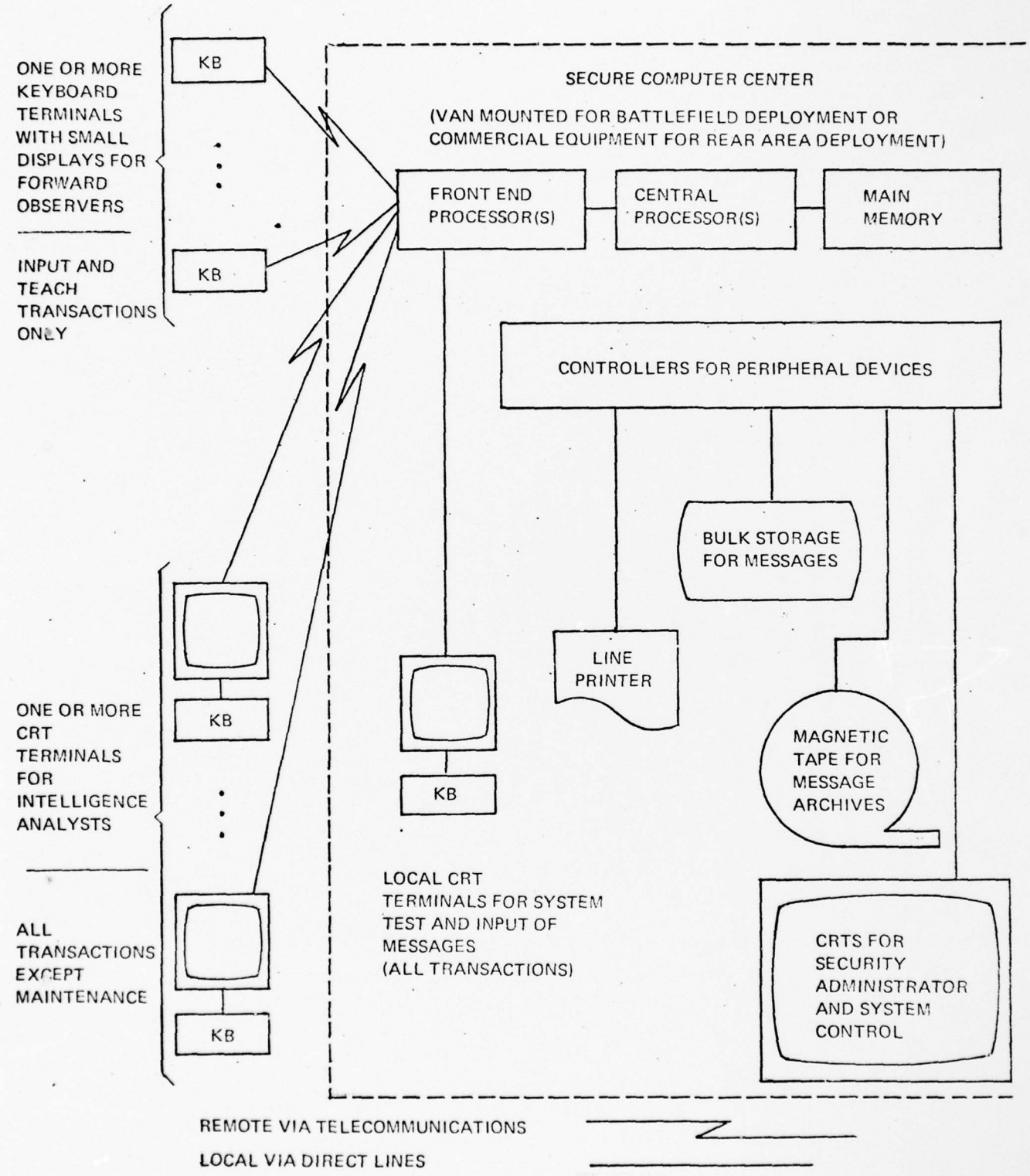


FIGURE 5-1. FUNCTIONAL EQUIPMENT/NETWORK DIAGRAM.

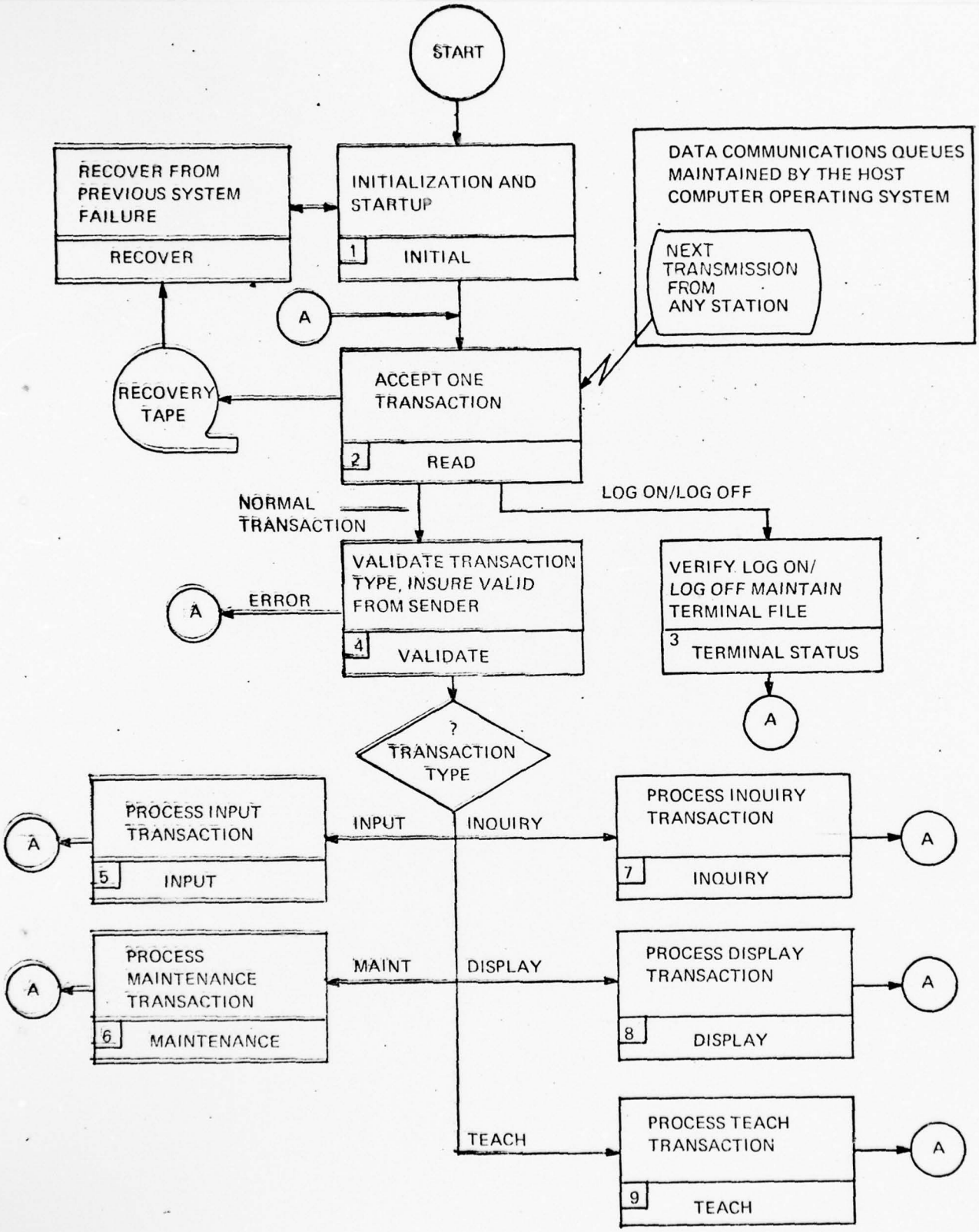


FIGURE 5-2. GLOBAL FLOW CHART

Table 5-3

File Composition

FILE NAME	CONTENTS	ONE RECORD FOR EACH
Terminal	terminal code on/off status last time/date used highest authorized security level location code list of authorized user codes by transaction type	terminal
Message	senders time/date senders ID receivers ID encoding indicator header message type message priority message reliability message security level message	
Log	A. Traffic Data no. of messages by terminal by transaction type during current (non archival) period B. Security Violation Data no. of violations by transaction type and by terminal C. Statistical Data no. of message by priority, security level no. of messages by transaction type by terminal by N no. of frequency distribution of me messages	Day
CLASSES	Names & Descriptions	Class
FACTORS	Names & Descriptions	Factor
TEACH		Transaction Type

Table 5-4

Transaction Types

TYPE	*PURPOSE*
INPUT	Submit of messages to the system
MAINTENANCE	Maintain the network includes: adding, deleting, changing security codes of terminals, displays of security violations, transfering old messages to archives
INQUIRY	Find, count, display the number or content of messages meeting specified search conditions
DISPLAY	Display classes, factors, message statistics, location, security, or traffic for terminals on request
TEACH	Inform user on procedures to be used in performing the types of system transactions

Functional Characteristics

The description of the functional characteristics is divided into two parts. The first part describes initial capabilities. The second part describes enhancements that can be added to the system as desired.

Initial Functional Capabilities

The system possesses three major functions:

1. message input
2. message storage
3. message and statistics retrieval

The additional requirements, SECURITY and RELIABILITY/RECOVERABILITY, are embedded in these major functions. A separate, but important, additional function is MAINTENANCE of the terminal network and system. Consider first the individual requirements of message input, storage, retrieval, system maintenance, and reliability separately. Security is discussed within each discussion.

Message Input

The system must allow many users to enter simultaneously complete messages including message text, headers, and the category(ies) in which each message is to be placed (this categorization could conceivably be performed automatically). While the average message is short (would fit on one CRT terminal screen), the system must be able to handle much longer messages. The system must be flexible and forgiving of user errors. It must be able to accept messages from all levels of experience. It must prompt the novice user for each piece of information required but allow the experienced user to enter all data without time wasting, extraneous communication. It must accept the elements of the header in any order. No message should be accepted by the system until the security of the user and the security of the terminal have been verified. Any security violations should be reported to the security officer immediately and logged.

Message Storage

The system must be able to store messages at two levels: immediate access and archival. Additionally, the system should allow groups of messages to be moved from one level to another, as described, without affecting the ability to store incoming messages.

Retrieval

The system must allow a small number of retrieval stations to access simultaneously stored messages in user-oriented conversational language. There are two levels of inquiries: those which ask for actual message text and those which ask for statistics about groups of messages.

Before allowing any retrieval transactions, the system must verify that both the station (terminal) and user are cleared for inquiry transactions for text or statistics as appropriate. In addition, before any message text can be transmitted to a station, the clearance of the station and user must be as high as the security level of the message.

The user must be able to form a set of messages selected by one or more:

a. category
b. class
c. factor
d. date
e. security
f. reliability
g. sender
h. addressee
i. key word or phrase

The user should be able to combine selected sets of messages and be able to form subsets of selected messages by further refining the selection criteria.

For any selected set of messages, the system should be able to display the message text or display statistics on the messages in the set. The available statistics must include simple message counts and tabulations or plots or other visual display of the relationship between pairs of selection variables.

Messages and statistics should be displayed on the user's terminal (CRT) and/or printed by a high speed printer.

A complete set of "TEACH" functions must be provided as an aid to the user.

Maintenance

Since the maintenance functions directly affect the security and reliability of the system, these functions must be performed only from within the physically secure confines of the main computer facility.

The maintenance functions should include at least:

a. add or delete a terminal from the network
b. modify the security level, functional capabilities, and valid users of a terminal
c. modify the user's security or capabilities
d. monitor the network, both in real time and by placing all traffic to and from a set of terminals into a log
e. force archival storage or recall of any selected set of messages

Reliability

Since the proposed system is not part of a tactical message transmission system, it does not need to be as reliable as a store-and-forward message processing system. The reliability levels of commercially available computer systems should be acceptable for present needs. However, the system must have recoverability features to minimize operator and user confusion after a system failure.

The functions that need to be performed after a system failure are:

a. force each user to log on before receiving or inserting any further data
b. if an entire transaction was received from a user (but was not completely processed), restart the processing for that transaction
c. if only part of a transaction was received from a user, show the user what was received and ask for the remainder
d. report any unrecoverable situation to operations

Enhancements

Two primary areas exist for additional enhancement and sophistication: the input phase and the inquiry phase.

In the area of message input, the system could recognize actual tactical message formats. This would allow message input to be essentially a clerical task and to be separated from the task of category assignment. Messages could then be input by clerk typists, optical scanners, paper tape, or a direct connection with a tactical automatic message processing system. Once a message has been received by the system, it could be displayed on a CRT terminal for actual analysis and category assignment. These enhancements are easily within current technology. The next enhancement could be to have the system itself read and analyze each message as received, and assign message categories. This enhancement possesses some technical risk for arbitrary messages.

The second area of possible enhancement is related to message inquiry. The class of possible questions could be made more general. The possibility exists that the system can be enhanced so that it will suggest an "answer" or suggested course of action relative to a query. Such a "decision recommending" feature would require research into the "decision policy" of current analysts. However, such "policy capturing" has been demonstrated to be entirely feasible (Slovic & Lichtenstein 1970; Dawes, 1970; Christal, 1963; Stephanson & Ward, 1971).

Similarly, the system could be enhanced to perform predictions and trend analysis. Predictions could be made relative to presumed enemy goals and/or tentative friendly actions. Answers could be based on a "reliability factor" to insure that only "best" answers are provided. In cases where content is contradictory, multiple answers could be provided with assigned probabilities.

The modular design of the proposed system permits the addition of any of these enhancements in an orderly fashion without affecting existing capabilities. This design also permits other kinds of functions to be added. Large size color CRT's could be added to show terrain and weather with past, current, or predicted future troop positions.

These suggested enhancements show possible system expansion. Some of these features can be readily implemented at this time. Others would require some effort for full implementation. All of these features can be performed on currently existing hardware, and most features have been demonstrated to some extent within the current software state of the art.

Program Implementation

The program would be implemented as a transaction processor system and is divided into three major components: (1) central transaction processor, (2) modules that process individual transactions, and (3) auxiliary routines.

The central transaction processor and global system flow are presented in Figure 5-2. The central transaction processor consists of four modules: INITIAL, READ, TERMINAL STATUS, and VALIDATE.

Module 1: INITIAL

The INITIAL module performs system initialization either at system startup or after a system failure. It initializes all data areas and checks for existing copies of files. In the event of a restart after a system failure, INITIAL calls for the RECOVER module (#22). INITIAL marks all terminals as active or inactive and logged off. When all initialization and recovery actions are complete, control is passed to the READ module.

Modules 2, 3, and 4 form a closed central loop. Controlled by this loop, but as separate tasks, are variable number of copies of the processing modules. Each copy of a processing module completely processes one request at a time. When completed, it checks for another request in its queue. If present, it processes that request. If its queue is empty, the module will wait for a short period and, if its queue remains empty, it will be discontinued. This feature frees system resources and helps to ensure that the system is always balanced for the current mix of requests.

Module 2: READ

The READ module accepts a transmission from the data communications queue for any station, writes the transmission to the recovery tape with date, time, and sender information, and then checks the transmission's content. If the transmission is a log-on or log-off request, the transmission is passed to the TERMINAL STATUS module (#3). If the transmission is not a log-on or log-off request, the terminal file is checked to ensure that the sending terminal is logged on. If not, a security error is reported and the message dropped. If the transmission is the first part of a new transaction, it is passed to the VALIDATE module (#4). If the transmission is a continuation of a current in process transaction, the transmission is queued to the particular copy of a processing module which has started processing that transaction.

Module 3: TERMINAL STATUS

The TERMINAL STATUS module handles all log-on and log-off activity. If a log-on request, the module verifies that the user is a valid user of that terminal. If not, a security error is reported, and all further activity from that terminal is ignored until the security officer takes appropriate action. In any event, the terminal file is updated to show the latest status of the terminal.

Module 4: VALIDATE

The VALIDATE module performs initial validation on transactions. It verifies that the specified transaction is valid for the specific user. Any errors are reported to the user and are logged. If the transaction is valid, it is queued for one of the processing modules on the basis of its type. If an idle copy of a program module of the required type is available, transaction is given to that copy. If no idle module copy is available, the transaction may be queued to a busy module to be processed when that module completes its current task, or a new copy of a module may be initiated. This decision is based on number of queued transactions, number of active processing modules, message priority, and expected and desired response times. Once the transaction has been queued to a processor module, control returns to the READ module (#2).

Module 5: INPUT

The INPUT module processes all message input. It verifies header, text, and category information. If no errors are detected, the message is stored in immediate message storage. Keys are built for specified categories and header information for quick access. If an error is detected on input, the user is lead to the correct response by a series of prompts. Successful storage of the message is reported to the user and logged (Figure 5-3).

Module 6: MAINTENANCE

The MAINTENANCE module validates maintenance requests and reports errors to the user by means of prompts (Figures 5-4). The validated request is processed, given to the user, and logged.

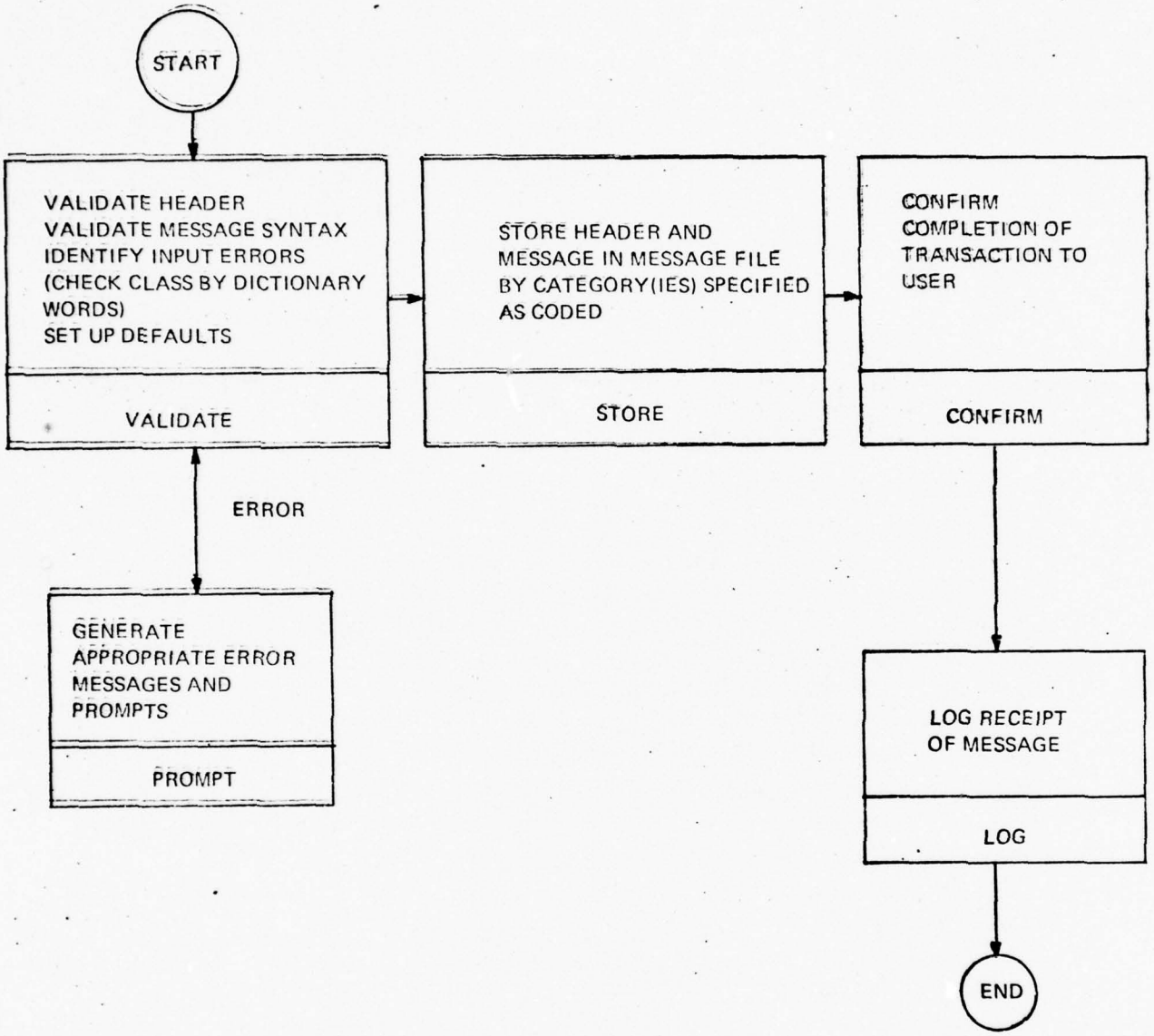


FIGURE 5-3. INPUT MODULE FLOW

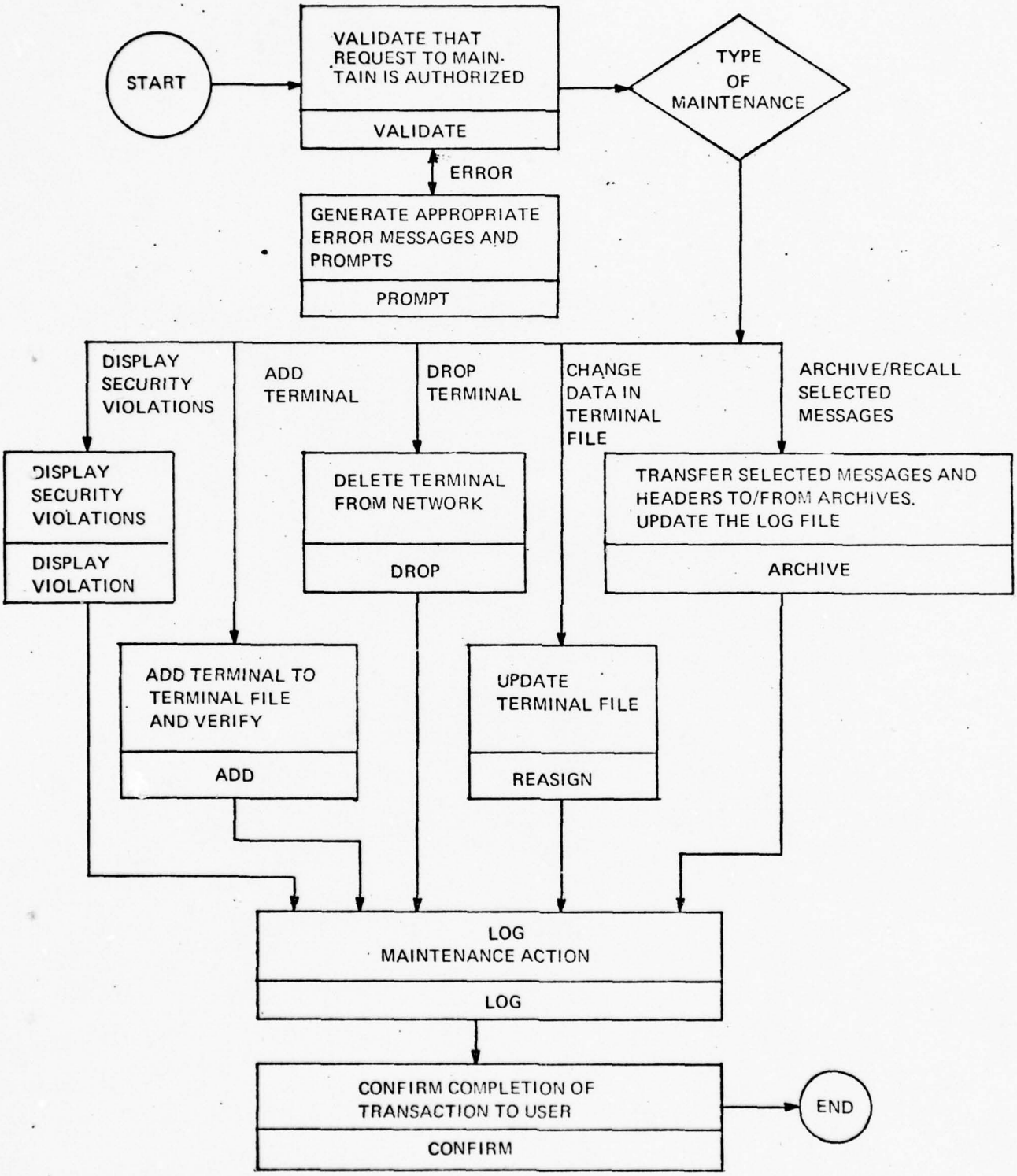


FIGURE 5-4. MAINTENANCE MODULE FLOW

Module 7: INQUIRY

The INQUIRY request module validates any errors corrected by means of prompts (Figure 5-5). The request is processed (this may involve sequential searching of a large number of messages), and the final formatted results are sent to the requester. If the request is expected to take a time which exceeds a predetermined threshold, the sender is notified that his request is in process. Before actual message text can be sent to the user, the security of the user and terminal are checked. The request is logged as completed when the user has received his requested information.

Module 8: DISPLAY

The DISPLAY module validates a request and any errors are corrected by means of prompts (Figure 5-6). The requested information is returned to the user, and the completion of the request is logged.

Module 9: TEACH

The TEACH module validates a request and the appropriate information is displayed to the user from the teach file (Figure 5-7). The request is logged when complete.

There are three auxiliary modules: ERROR, LOG, and RECOVERY. These are used for special functions by other modules.

ERROR

The ERROR module is used to report syntax errors to the user and security errors to the security officer. All errors are logged for later performance evaluation. (This module is not shown in the flow logic drawing since it may be called from any PROMPT or VALIDATE module.)

LOG

The LOG module is responsible for adding information to the log file. A history of all action taken by the system is contained in the log. The receipt of each request and its completion is logged along with date, time, and user identification. All errors and maintenance actions are also logged, as are all recovery actions. By appropriate analysis of the log, questions about system or user performance can be answered.

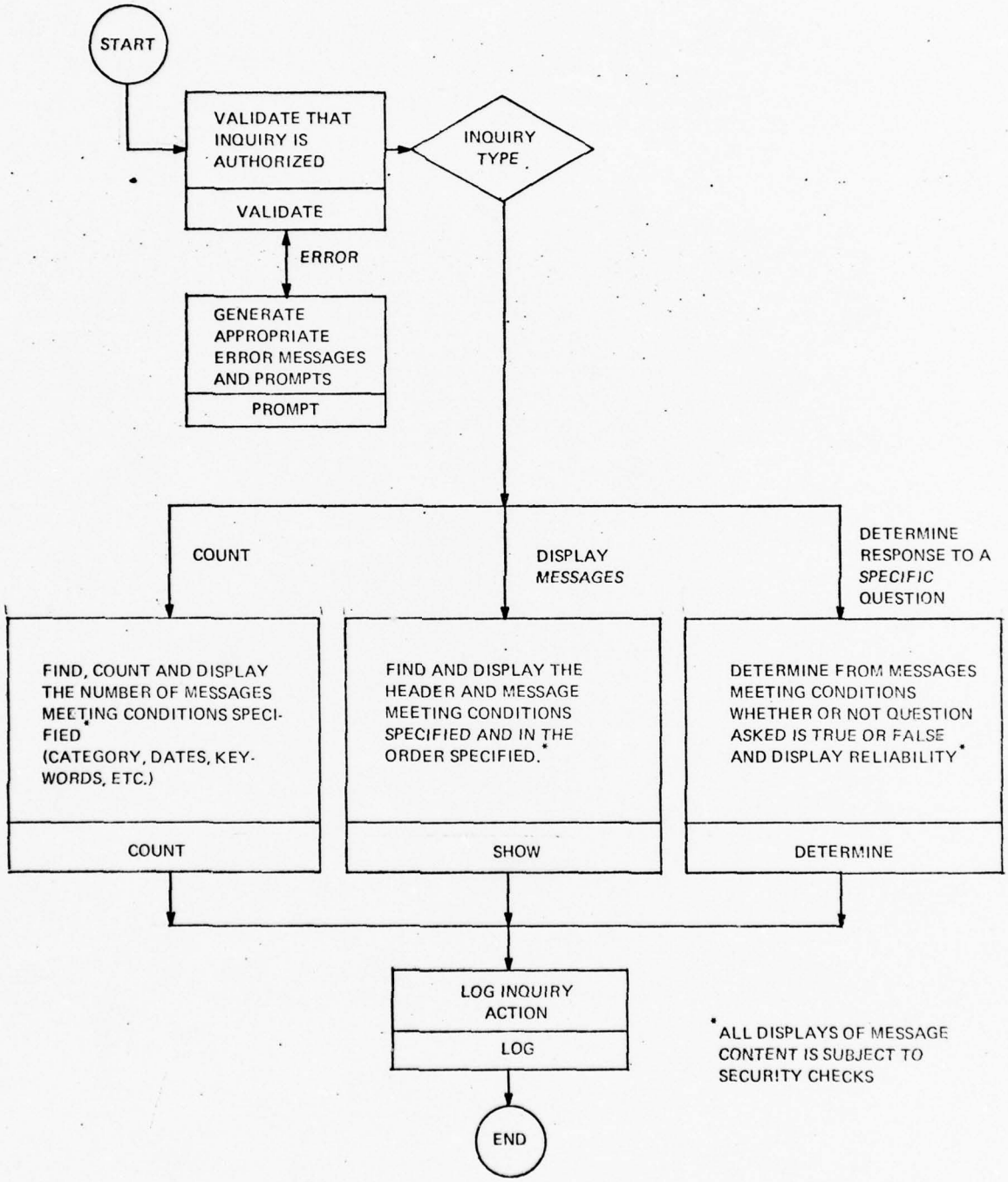


FIGURE 5-5. MESSAGE CACHE INQUIRY FLOW

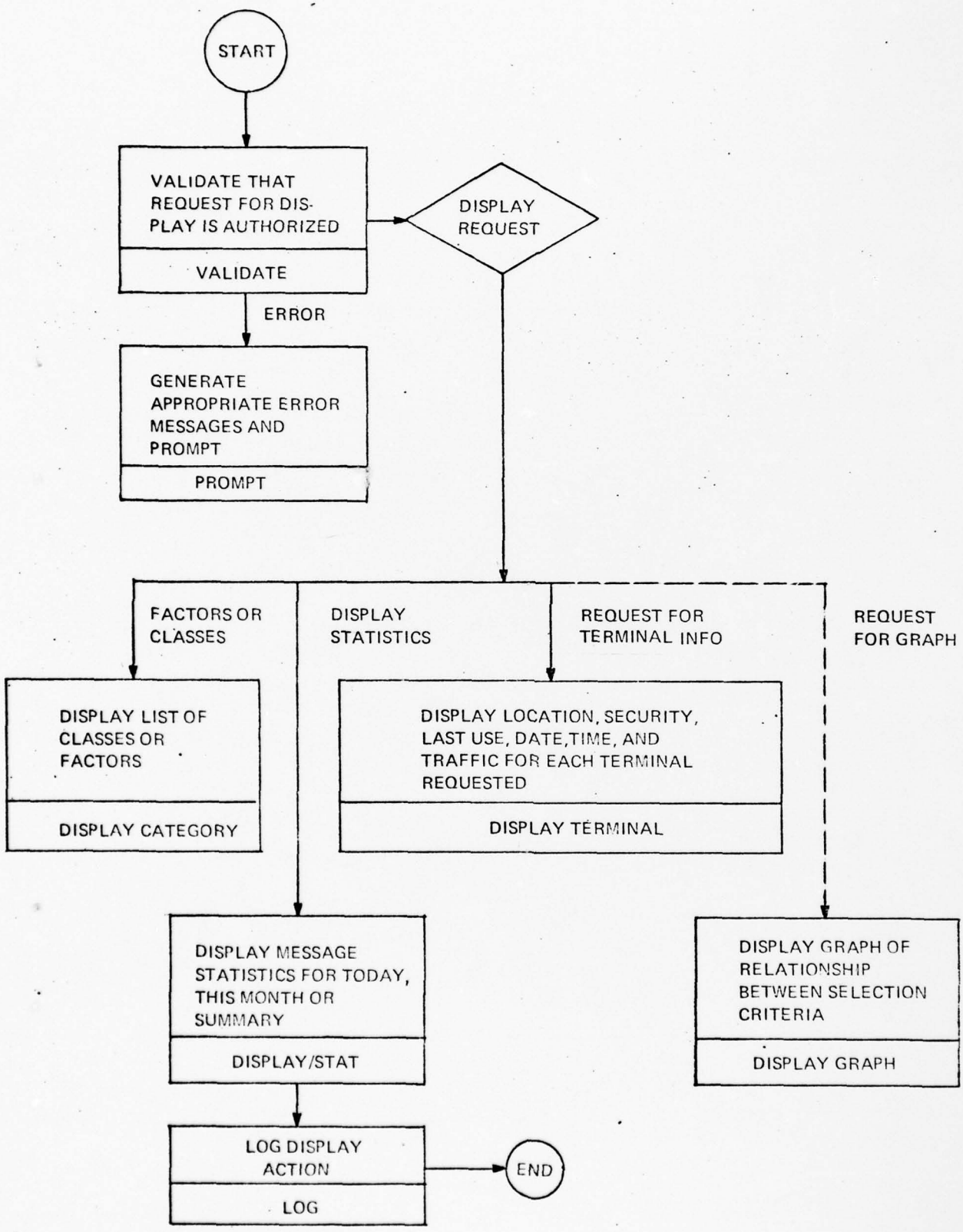


FIGURE 5-6. DISPLAY MODULE FLOW

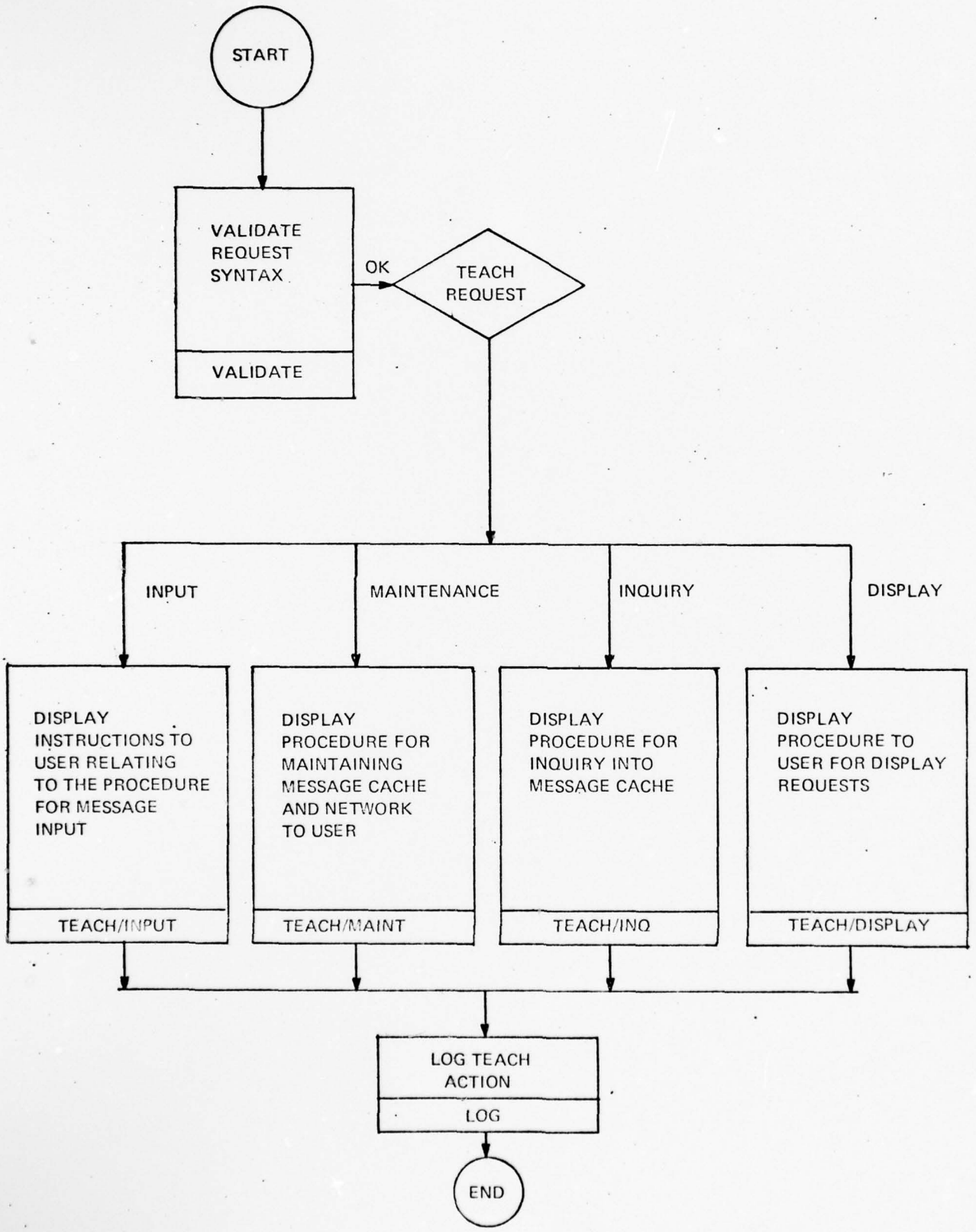


FIGURE 5-7. TEACH MODULE FLOW

RECOVERY

The RECOVERY module is called only after a system failure. By reading the recovery tape and matching each transaction on the tape with the log, the status of each transaction at the time of system failure can be determined and appropriate action taken.

CHAPTER VI

DISCUSSION AND CONCLUSIONS

The present study was instituted to determine whether or not a natural order could be derived for battlefield language. Such a natural order was, in fact, shown to be possible. It was possible to derive a model which is based on two complementary logics.

The first logic (multidimensional scaling) was based on derivation of the factorial structure of the perceptions of battlefield analysts of battlefield language. The isolated factors form one axis of the final model. This axis, which might be called the perceptual axis, is rooted in the language perceptions of battlefield analysts.

The second logic was based on application of concepts of transformational grammar. The results of this analysis yielded the second axis (classes) of the final model. This second axis might be called a linguistic axis. The linguistic axis is based on the structural aspects of battlefield language.

Accordingly, it can be argued that the final model is compatible with both the actual battlefield language structure and the perceptions of battlefield language of intelligence analysis. As suggested by Halpin (1975), use of such a model in a computer oriented message management system would make the equipment system compatible with the language and the perceptions of the user. Such an approach seems entirely rational from the integrated man-equipment system point of view. Within such a design construct, the information management system is based on the language and perceptions of the user rather than the reverse situation in which the user must adapt to machine requirements. Unfortunately, this latter design approach is the one found in most information management systems.

It is possible also that the high obtained useability (reliability) of the derived taxonomy, as evidenced by the two verificational experiments, is a result of the naturality built into the taxonomy. In view of the only limited training given to the participants in the two verificational experiments, the useability indications were higher than might have been anticipated. While other explanations of this finding are possible, the naturality concept represents an attractive explanatory construct and one which supports the utility of the final model.

PRECEDING PAGE BLANK-NOT FILMED

The high obtained degree of congruency between the factorial structure of the four military intelligence analysts whose perceptions formed the basis for the four separate multidimensional scaling analyses might also have been anticipated. The meaning and interpretation of words, phrases, and sentences is learned. This holds whether a total language or a sublanguage is involved. The military/battlefield subculture may serve to crystalize the battlefield language perceptions just as any other subculture fosters common language percepts within the subculture's members. For example, the sentence "He is spaced out" would probably be commonly perceived by an American teenage population but may be perceived differently by Russian teenagers. Unfortunately, the present study did not include a nonmilitary control group. Factorization of the percepts of control analysts of the same stimulus message set would be of interest relative to the homogenization conjecture.

Certainly other tests of the utility and veridicality of the taxonomy are indicated. The present tests, as reported in Chapter IV, were based on abbreviations of the total taxonomy in its form at that time. The question of whether or not the results would still obtain when the total taxonomy is employed remains open. However, we note that the final taxonomy represents considerable consolidation to reduce ambiguity. Such ambiguity reduction would be anticipated to increase utility. Accordingly, there are grounds for believing that experimental verification of the final model would yield results which are largely in conformity with the two verificational studies which were completed.

Finally, it seems that the battlefield information processing and intelligence derivation technology will eventually move towards a computerized approach such as that described in Chapter V. Such a system would lead towards the ability to perform online historical analysis, derivation of most probable enemy intent, trend analysis and prediction, and computer suggested course of action.

Conclusions

In view of the data presented in earlier chapters of the present report, the following conclusions seem warranted:

1. For the intelligence analyst sample (and battlefield messages involved), the perceptual structure may be described by 15 orthogonal factors. These factors seemed common across the four sample members for whom separate multidimensional scaling analyses were performed.

2. From the linguistic point of view, the structure may be described by a limited set of classes.

3. The combination of perceptual factors and linguistic classes into a taxonomy yielded a product which appears to be useful from both the information categorization and the information retrieval points of view. Analysts were able to classify/retrieve messages relative to the taxonomy, and analysts were sensitive to and confident in the quality of their performance when employing the taxonomy.

4. It appears as if a reasonable start has been made toward establishing how battlefield messages have attributes in common and attributes which are different.

REFERENCES

Abelson, R.P. A technique and a model for multidimensional scaling. *Public Opinion Quarterly*, 1954, *18*, 405-418.

Andrews, T.G., & Ray, W.S. Multidimensional psychophysics: A method for perceptual analysis. *Journal of Psychology*, 1957, *44*, 133-144.

Baker, J.D. Quantitative modeling of human performance in information systems. *Ergonomics*, 1970, *13*, 654-664.

Brown, K.R. Job analysis by multidimensional scaling. *Journal of Applied Psychology*, 1967, *51*, 469-475.

Butzkamin, J., & Pfeiffer, M.G. *Convergent and discriminant validity of college-level teaching dimensions.* Philadelphia, LaSalle College, 1973.

Campbell, D.T., & Fiske, D.W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 1959, *56*, 81-105.

Christal, R.E. *JAN: A technique for analyzing group judgment.* Lackland AFB, Tex., Personnel Research Laboratory and Aerospace Medical Division, PRL-TRD-63-3, February 1963.

Coombs, C.H. *A theory of data.* New York: Wiley, 1964.

Coombs, C.H., Dawes, R.M., & Tversky, A. *Mathematical psychology: An elementary introduction.* Englewood Cliffs, N.J.: Prentice-Hall, 1970.

Dawes, R.M. A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist*, 1971, *26*, 180-188.

Dixon, W.J. (Ed.), *BMD: Biomedical Computer Programs.* Berkley: University of California Press, 1973.

Egan, J.P., & Clarke, F.R. Psychophysics and signal detection. In J.B. Sidowski (Ed.), *Experimental methods and instrumentation in psychology.* New York: McGraw-Hill, 1966.

English, H.B., & English, A.C. *A comprehensive dictionary of psychological and psychoanalytical terms.* New York: Longmans Greene, 1958.

PRECEDING PAGE BLANK-NOT FILMED

Ekman, G. Eine neue Methode zur Erlebnisanalyses. *Zietschrift fuer Experimentelle und Angewandte Psychologie*, 1954, *2*, 167-174.

Ekman, G., & Sjoeberg, L. Scaling. *Annual Review of Psychology*, 1965, *16*, 452-474.

Ekman, G. Two methods for the analysis of perceptual dimensionality. *Perceptual and Motor Skills*, 1965, *20*, 557-572.

Ekman, G. Comparative studies on multidimensional scaling and related techniques. In K. Pawlik (Ed.), *Perspectives in multivariate psychological research*. Bern and Stuttgart: Huber, 1969.

Ekman, G. Comparative studies on multidimensional scaling and related techniques. *Reports from the Psychological Laboratories, University of Stockholm*, 1970, Suppl. 3.

Finley, D., Muckler, F.A., Gainer, C.A., & Obermayer, R.W. *An analysis and evaluation methodology for command and control: Final technical report*. Northridge, Calif.: Manned Systems Sciences, 1975.

Fleiss, J.L. Measuring nominal scale agreement among raters. *Psychological Bulletin*, 1971, *76*, 378-382.

Green, D.M., & Swets, J.A. *Signal detection theory and psychophysics*. New York: John Wiley, 1966.

Green, P.E., & Carmone, F.J. *Multidimensional scaling and related techniques in marketing analysis*. Boston: Allyn and Bacon, 1970.

Gross, M. *Transformational analysis of French verbal constructions*. Transformations and discourse analysis papers 74. Philadelphia: University of Pennsylvania, 1966.

Halpin, S. Personal communication, 1975.

Harris, Z. *Mathematical structures of language*. New York: Interscience, 1968.

Hiz, D., & Joshi, A.K. Transformational decomposition. *Proceedings of IFIP International Conference on Computational Linguistics*. Grenoble: 1968.

Jackson, D.M., Messick, S.J., & Solley, C.M. A multidimensional scaling approach to the perception of personality. *Journal of Psychology*, 1957, *44*, 311-318.

Johnson, S.C. Hierarchical clustering schemes. *Psychometrika*, 1957, *32*, 241-254.

Kaiser, H.F., Hunka, S., & Bianchini, J.C. Relating factors between studies based upon different individuals. *Multivariate Behavioral Research*, 1971, *6*, 409-422.

Kuennapas, T. Visual perception of capital letters: Multidimensional scaling and multidimensional similarity. *Scandanavian Journal of Psychology*, 1966, *7*, 189-196.

Kuennapas, T. Visual memory of capital letters: Multidimensional ratio scaling and multidimensional similarity. *Perceptual and Motor Skills*, 1967, *25*, 345-350.

MacCallum, R.C. Relations between factor analysis and multidimensional scaling. *Psychological Bulletin*, 1974, *81*, 505-516.

Messick, S.J. *The perception of attitude relationships. A multidimensional approach to the structuring of social attitudes.* Unpublished doctoral dissertation. Princeton University, 1954.

Messick, S.J. An empirical evaluation of multidimensional successive intervals. *Psychometrika*, 1956, *21*, 367-375.

Miles, J. Eliminating the guessing factor in the multiple choice test. *Educational and Psychological Measurement*, 1973, *33*, 637-651.

Miller, R.B. Task Taxonomy: Science or technology? In W.B. Singleton, R.S. Easterby, & D. Whitfield (Eds.), *The human operator in complex systems*. London: Taylor and Francis, 1967.

Miron, M.S., Patten, S.M., & Halpin, S.M. *Context analysis and the organization of combat intelligence data*. Syracuse University Research Corporation and US Army Research Institute for the Behavioral and Social Sciences, 1975.

Pfeiffer, M.G., Kuennapas, T., Fastiggi, C.F. Common elements approach to multidimensional similarity analysis among the job tasks. *Perceptual and Motor Skills*, 1973, *36*, 3-12.

Reeb, M. How people see jobs. A multidimensional scaling analysis. *Occupational Psychology*, 1959, *35*, 659-660.

Richardson, M.W. Multidimensional psychophysics. *Psychological Bulletin*, 1938, *35*, 659-660.

Schultz, C.L., Brumback, G.B., & Rizzo, J.R. An approach to dimensions of value. *Journal of Applied Psychology*, 1964, *57*, 101-201.

Sells, S.E. (Ed.) *Stimulus determiantion of behavior*. New York: Ronald, 1963.

Siegel, A.I., & Pfeiffer, M.G. Predicting academic success through application of theory of signal detectability variables. *Proceedings of the 77th. Annual Convention of the American Psychological Association*, 1969, *4*, 145-146.

Siegel, A.I., & Bergman, B.A. *Adaptation of advanced measurement and evaluation techniques for utilization in Air Force Technical Training Systems*. Lowry AFB, Colo.: Air Force Human Resources Laboratory, 1973.

Silver, C.A., Landis, D., & Jones, J.M. *Development of criteria for evaluation of large-screen displays*. Rome, NY: RADC-TR-65-198, 1965.

Sjoberg, L. Models of similarity and intensity. *Psychological Bulletin*, 1975, *82*, 191-206.

Slovic, P., & Lichtenstein, S. Analyzing the expert judge: A descriptive study of a stockbroker's decision process. *Journal of Applied Psychology*, 1969, *53*, 255-263.

Smith, R., & Siegel, A.I. A multidimensional scaling analysis of the job of civil defense director. *Journal of Applied Psychology*, 1967, *51*, 476-480.

Stephanson, R.W., & Ward, J.H. Computer assisted discussions to help a policy group assign weights to criterion ratings. *Personnel Psychology*, 1971, *24*, 447-461.

Stone, L.A., Coles, G.J., & Lindem, A.C. *Multidimensional evaluation structure analysis*. Grand Forks, Indiana: Judgmetrics, 1970.

Stone, L.A., & Coles, G.J. Dimensions of color vision revisited. *Journal of Applied Psychology*, 1971, *77*, 79-87.

Swets, J.A. (Ed.) *Signal detection and recognition by human observers*. New York: John Wiley, 1964.

Wright, A.A. Psychometric and psychophysical theory within a framework of response bias. *Psychological Review*, 1974, *81*, 322-347.

APPENDIX A

Loading of Each Stimulus on Each Dimension--
Final Factor Analysis

Dimension

	I	II	III	IV	V	VI	VII	VIII	IX	X
VARIABLE 1										
	-0.29243	0.04874	0.04738	0.05898	0.12574	-0.32952	0.19291	-0.50436	-0.15229	0.14628
	0.40230	0.02531	0.31770	0.09780	-0.18160	0.04096				
VARIABLE 2										
	-0.18122	-0.03694	-0.13552	-0.03950	-0.07752	-0.86784	0.00521	-0.16017	-0.15261	0.07756
	0.13605	0.08810	0.05380	0.03634	-0.03517	-0.01985				
VARIABLE 3										
	-0.53755	-0.00742	0.21655	0.30603	-0.04213	-0.38967	0.08274	-0.12660	-0.08644	0.20678
	0.13835	0.09198	0.45160	0.02082	0.07747	0.05921				
VARIABLE 4										
	-0.00607	-0.14543	-0.81090	-0.04646	0.12463	-0.20253	0.12549	0.12896	0.01806	-0.12355
	-0.02917	0.01877	0.01109	0.02217	0.06131	0.01740				
VARIABLE 5										
	-0.59461	0.16811	-0.47813	0.27862	-0.10980	0.03403	-0.11333	-0.09700	-0.15135	0.05592
	-0.02631	-0.17145	-0.09892	0.20030	0.20284	-0.03816				
VARIABLE 6										
	-0.56557	-0.05719	-0.61603	0.14224	-0.04518	-0.21819	-0.14039	-0.01776	0.08805	-0.00117
	0.18929	-0.05626	-0.07572	-0.02971	0.14009	0.05894				
VARIABLE 7										
	-0.06706	-0.09855	-0.21420	0.05985	-0.08057	-0.84541	0.01518	-0.03772	-0.28406	0.07076
	-0.01919	-0.02049	0.09638	0.15429	-0.03763	0.00431				

PRECEDING PAGE BLANK-NOT FILMED

BEST AVAILABLE COPY

VARIABLE 8									
-0,07885	0,12505	-0,03381	-0,04491	-0,06099	0,00376	0,92543	-0,10574	+0,06599	+0,01793
0,04272	0,05407	0,14057	-0,04278	0,03201	0,04783				
VARIABLE 9									
-0,09938	0,05612	-0,88474	0,07706	0,01939	-0,10889	+0,08030	-0,03611	-0,00064	0,12237
-0,02233	-0,05264	-0,03036	0,03466	0,06154	0,03679				
VARIABLE 10									
0,06714	-0,00084	-0,76921	0,10192	-0,05182	-0,12440	0,24335	0,00036	0,12935	0,26204
-0,00281	-0,15398	-0,02205	0,03231	-0,04765	-0,28841				
VARIABLE 11									
0,01322	-0,03933	0,02770	-0,00597	0,00562	0,03770	0,89120	+0,00165	+0,15389	+0,12374
0,11032	0,05031	-0,12030	0,04927	-0,13454	-0,01461				
VARIABLE 12									
-0,90339	0,06359	-0,00900	0,04444	-0,02129	-0,21586	-0,05269	+0,01996	+0,05670	+0,06079
-0,08547	0,03201	-0,00122	-0,07379	-0,15290	-0,01766				
VARIABLE 13									
0,11948	-0,16258	-0,12244	-0,29771	-0,30203	-0,25951	0,03180	-0,24623	+0,11060	0,33288
-0,01567	0,13836	0,02206	0,18216	0,07939	0,49405				
VARIABLE 14									
-0,00566	0,19569	-0,15905	-0,09213	-0,03079	+0,17522	0,32866	-0,45883	+0,41287	+0,05897
-0,18969	-0,25075	0,06560	0,17791	-0,28010	+0,15351				
VARIABLE 15									
-0,23001	-0,06639	0,01197	0,25709	-0,72381	-0,29900	0,02998	0,10562	+0,04693	0,15868
0,08836	-0,07931	0,24057	0,03141	0,05723	0,09107				
VARIABLE 16									
-0,09070	0,47070	0,35020	0,04595	0,09432	-0,04875	0,11948	-0,23465	+0,12475	+0,09727
-0,03155	-0,00038	0,00790	0,01879	-0,63869	0,00156				
VARIABLE 17									
-0,74434	-0,03082	-0,23610	-0,03719	0,15928	0,02550	0,09658	0,06916	+0,33391	+0,10019
0,05224	0,12663	0,13415	0,08819	0,04833	0,03479				
VARIABLE 18									
0,06400	-0,04609	0,22669	-0,84762	-0,16008	0,07208	0,02703	-0,08441	+0,09704	+0,16601
0,01988	-0,04708	-0,01831	0,00049	-0,04869	+0,10299				
VARIABLE 19									
-0,16279	0,05944	0,02167	-0,28095	0,06908	+0,09983	0,07180	-0,77516	+0,14130	+0,08895
0,15185	0,03530	0,05683	-0,09347	-0,26810	+0,05183				
VARIABLE 20									
0,21562	-0,13219	0,26036	-0,41503	-0,40935	0,20508	0,04816	-0,22055	0,11525	+0,04738
0,19701	0,21977	-0,12693	0,35162	-0,13820	0,10628				
VARIABLE 21									
0,01568	-0,17928	0,08374	-0,17578	-0,24342	-0,10144	0,10426	-0,66301	0,02151	0,15471
0,20477	0,31251	0,10780	0,24854	-0,00252	0,20791				
VARIABLE 22									
-0,17306	0,08551	-0,20132	0,25708	-0,01660	-0,24961	-0,07445	-0,72813	+0,04200	+0,06343
-0,24720	-0,21732	-0,05299	0,02976	0,00484	-0,00158				
VARIABLE 23									

BEST AVAILABLE COPY

0.08205	-0.01041	0.13791	-0.82322	-0.05340	0.05754	0.08295	-0.10968	0.11204	0.03371
-0.01744	0.13297	-0.03236	0.16201	0.02476	0.16683				
VARIABLE 24									
-0.67007	0.10949	-0.19861	-0.10682	0.12550	0.01155	-0.00360	-0.46197	-0.09409	-0.07184
-0.12671	-0.13696	-0.16846	-0.14902	0.04408	0.01256				
VARIABLE 25									
-0.10907	0.04575	-0.86644	0.04130	-0.00723	0.03740	-0.01943	-0.13548	-0.03201	-0.11241
-0.08215	0.01770	-0.01758	-0.11773	0.14878	0.06321				
VARIABLE 26									
-0.59022	0.03494	0.06377	0.33535	-0.35383	-0.14652	0.01602	-0.11441	-0.12909	0.13975
-0.06615	0.12691	0.42505	-0.01161	0.05272	-0.03621				
VARIABLE 27									
0.14336	-0.13813	-0.02518	-0.19535	-0.06546	-0.02040	0.02964	-0.01130	-0.12962	0.14814
0.09631	0.07919	-0.08833	0.05452	0.07244	-0.00894				
VARIABLE 28									
-0.04726	-0.09332	-0.00499	-0.14731	-0.05341	-0.18643	-0.01835	-0.10106	-0.44081	0.10303
-0.05540	0.16814	0.00211	0.74318	0.02065	0.05344				
VARIABLE 29									
-0.83512	0.01475	-0.10029	0.03986	-0.02970	-0.40167	-0.06249	-0.02910	-0.02400	0.07808
-0.03641	-0.06907	0.08049	0.13591	-0.06941	-0.04588				
VARIABLE 30									
-0.14935	-0.09299	-0.20927	-0.07102	-0.05092	-0.76071	-0.00281	-0.05296	-0.30529	0.10256
0.10506	0.24382	-0.06350	-0.04924	-0.08255	-0.05824				
VARIABLE 31									
-0.55244	0.20738	0.03188	-0.05265	0.01759	-0.57725	-0.03471	-0.32401	0.04498	-0.04499
-0.13792	0.06433	-0.15191	-0.07597	-0.02856	0.05377				
VARIABLE 32									
-0.69311	0.10289	-0.01862	-0.22211	0.09512	0.05864	0.06657	-0.55653	0.06791	0.05129
-0.02417	-0.01873	-0.16123	0.03164	0.02412	-0.04894				
VARIABLE 33									
-0.10261	0.42470	-0.10204	-0.10064	-0.19235	-0.11435	0.11999	-0.07182	0.21579	-0.15609
0.04096	0.53339	-0.22398	0.32659	0.25243	-0.19215				
VARIABLE 34									
-0.03755	0.12101	-0.18266	-0.71584	0.25527	-0.28024	-0.08166	0.00298	-0.09944	0.17904
-0.08429	0.02577	-0.03793	-0.25929	0.06397	0.02625				
VARIABLE 35									
-0.34970	-0.21772	-0.07034	-0.14229	0.13235	-0.09401	0.20806	-0.33824	-0.49861	0.09724
-0.04963	0.00137	-0.17415	-0.37425	-0.10817	-0.00663				
VARIABLE 36									
-0.33274	-0.00504	-0.00826	-0.00329	0.15020	-0.06943	0.16655	-0.14607	-0.73323	-0.14834
-0.19235	0.12756	0.16458	-0.03484	-0.05725	-0.06000				
VARIABLE 37									
-0.17946	-0.20009	-0.45453	0.38315	-0.21666	-0.37604	-0.19688	-0.31515	0.05728	0.08755
-0.06129	0.00096	0.11244	0.13633	0.16576	0.10100				
VARIABLE 38									
-0.09807	-0.09411	-0.32422	0.17271	-0.33639	-0.39243	-0.19817	-0.28826	0.14152	0.36564

BEST AVAILABLE COPY

-0,09344	-0,00830	0,11535	-0,01987	0,25120	0,20910				
VARIABLE 39									
0,19257	-0,36878	0,28404	-0,12067	-0,14953	-0,25788	0,07527	0,11201	-0,27617	0,21199
0,54849	-0,04159	-0,05099	-0,14298	0,05652	0,10016				
VARIABLE 40									
0,15099	-0,53669	-0,19263	-0,01814	-0,25253	-0,20188	0,13260	0,10220	-0,00039	0,11654
0,38800	-0,19378	-0,06153	0,40841	0,15709	0,03146				
VARIABLE 41									
0,01015	-0,37766	-0,12755	0,03783	-0,24584	-0,05892	-0,08840	-0,25295	-0,60963	0,29467
0,27820	-0,03760	-0,05030	-0,00781	0,20443	-0,17022				
VARIABLE 42									
-0,04019	-0,31930	-0,14761	0,22069	-0,04401	-0,04629	0,13922	0,05983	-0,54599	0,20596
0,32443	-0,22222	-0,11933	0,33578	0,00171	-0,18462				
VARIABLE 43									
-0,12030	-0,41415	0,13802	-0,06203	-0,00844	0,05449	0,06880	0,13039	-0,53900	0,00224
0,10427	-0,01220	0,03346	-0,05688	-0,40294	0,39600				
VARIABLE 44									
-0,32499	-0,66715	0,07137	0,16469	-0,28839	-0,13658	-0,05820	0,08068	-0,29465	0,11524
0,20326	-0,03454	0,11370	-0,05719	0,01433	-0,03487				
VARIABLE 45									
-0,01336	-0,76829	-0,11783	-0,12935	0,13546	-0,15523	-0,10811	0,06564	-0,04664	-0,07898
0,15650	0,02174	0,13906	-0,03394	0,12995	0,07985				
VARIABLE 46									
0,01396	-0,00217	0,28928	-0,01232	0,13523	-0,10576	0,04899	-0,22457	-0,37390	0,05422
-0,00703	0,09780	-0,00535	-0,09182	-0,74815	-0,05611				
VARIABLE 47									
-0,02912	-0,30623	0,03287	-0,01258	-0,27940	-0,00781	0,01208	0,03574	-0,22721	0,72217
0,09562	0,18405	-0,17197	0,03628	-0,16679	-0,10792				
VARIABLE 48									
-0,05742	-0,59788	-0,39579	0,17445	0,04738	0,00661	-0,17330	-0,11050	0,00397	0,10113
0,15160	-0,10621	-0,00454	-0,01389	-0,10714	-0,36526				
VARIABLE 49									
0,01884	-0,05508	0,10722	-0,06776	-0,02249	-0,27425	0,08080	0,01724	-0,11794	0,22988
-0,03699	0,78519	0,11205	0,01683	-0,11514	0,07058				
VARIABLE 50									
-0,09362	-0,10605	-0,10437	0,04704	-0,09519	-0,21686	-0,19431	0,01152	-0,07000	0,82444
0,12690	0,06318	0,19437	0,05009	0,10328	0,11159				
VARIABLE 51									
0,16483	-0,59045	0,06679	0,13576	-0,32147	-0,17952	0,26218	0,00964	-0,00909	0,26199
0,02027	-0,15135	-0,23433	0,29117	0,18900	0,06061				
VARIABLE 52									
0,07034	-0,23368	0,10905	-0,03761	-0,10052	-0,32660	0,05531	0,03131	-0,70595	0,11016
0,15822	0,10171	-0,01018	-0,01128	-0,13056	0,27565				
VARIABLE 53									
-0,15284	-0,35143	-0,35352	0,04110	-0,03213	-0,05509	-0,04310	-0,35942	-0,25564	0,01679
-0,11875	-0,47103	0,26717	-0,13293	-0,03913	-0,05011				

BEST AVAILABLE COPY

VARIABLE 54									
0,01021	-0,17196	0,09434	-0,07812	-0,15300	-0,09199	0,17059	-0,01966	-0,61145	0,00698
0,31972	-0,13846	0,00401	0,23306	-0,41823	0,24623				
VARIABLE 55									
-0,13749	0,02298	0,01690	0,05138	0,05001	-0,35756	-0,08252	-0,03688	-0,73471	-0,10051
0,13888	0,04289	0,23927	0,26660	-0,00804	-0,05656				
VARIABLE 56									
0,03076	-0,31045	0,01658	0,05231	-0,11376	-0,04700	0,12460	-0,03934	-0,17139	0,06940
0,79072	0,05283	0,02125	0,02828	-0,03043	-0,03318				
VARIABLE 57									
-0,63470	-0,31802	-0,18224	0,10086	0,12549	0,16024	0,23651	-0,00061	-0,10655	0,15871
0,28470	-0,07918	0,02946	-0,06945	0,12495	-0,03500				
VARIABLE 58									
0,16586	-0,76146	0,15133	-0,05347	-0,18942	0,14668	-0,00077	-0,03637	-0,13569	0,20048
0,06459	0,13802	-0,12804	0,11487	-0,06373	0,11485				
VARIABLE 59									
-0,01926	0,09326	0,19438	-0,09810	-0,21585	0,03745	0,07628	0,06458	-0,76374	0,27468
0,07050	0,02117	-0,06545	-0,01605	-0,10074	-0,19917				
VARIABLE 60									
-0,65619	-0,16195	0,29135	0,20752	-0,23509	-0,01632	0,08473	-0,07440	-0,14998	0,08473
-0,05698	-0,08753	0,40532	-0,17185	-0,21319	0,06441				
VARIABLE 61									
-0,19921	-0,11499	-0,04651	0,15720	0,04128	-0,38091	0,04990	-0,15248	-0,70289	0,00322
0,02438	-0,15445	-0,04396	0,07961	-0,11796	0,26027				
VARIABLE 62									
-0,09751	-0,23833	-0,52094	0,32514	-0,16943	-0,45010	-0,22792	-0,16748	-0,04592	0,03294
0,06457	0,07459	0,03224	0,13327	0,22238	-0,04737				

BEST AVAILABLE COPY

APPENDIX B

Dictionary of Battlefield Language Based on Derived Taxonomy

PRECEDING PAGE BLANK-NOT FILMED

I. INTRODUCTION

This dictionary describes a preliminary system for classifying battelfield language. Basically, the system involves technique for classifying battlefield messages into categories according to two classification schemes. These two schemes each partition the whole range of battlefield language. One scheme consists of 15 divisions; each of these is called a "factor." The second scheme consists of 27 divisions; each of these is called a "class."

Within the system, a battlefield message may be classified using both classes and factors in combination (called categories). Retrieval of information from this system is accomplished by selecting for review all information in one or several categories.

This Dictionary provides all the basic information required to use this system. Sections II and III list and define the factors and classes. Section IV consists of a model implementation of the system. This model serves as a reference which may be consulted if the use of any category is unclear. For the purposes of showing further use of the system, two analyses of the phrases shown in the model are presented in sections V and VI. Section V alerts the user to synonyms in the battlefield language. Section VI contains an alphabetical list of all the words used in the model and defines these words in terms of categories in which they are found.

PRECEDING PAGE BLANK-NOT FILMED

II. LIST OF FACTOR AND CLASS NAMES

Factor Names

A. Movement and Capability for Movement

B. Numbers, Trends, and/or Source

C. Type of Activity

D. Training, Morale, and Attitudes

E. Communications

F. Armament Preparation

G. Small Arms Capability

H. Strength and/or Location of Units

I. Plans and Expectations

J. Extra Military Information Sources

K. Deserter Information

L. Logistics

M. Unique Actions

N. Contents of Documents

O. Civilian-Military Relationships

Class Names

1. Composition of Units
2. Markings and Insignia
3. Position of Forces and Equipment
4. Strength of Units
5. Replacements
6. Quantitative Indications
7. Weapons and Equipment
8. Analysis
9. Characterization of Missions
10. Sabotage
11. Actual and Threatened Contact
12. Strategy and Tactical Doctrines
13. Preparations and Readiness
14. Movement
15. Chronology
16. Information about Officers
17. Competence of Personnel and Equipment
18. Mental and Physical Condition of Troops
19. Civilian-Military Relations
20. Sources/Methods of Obtaining Information
21. Informer Source-Military
22. Informer Source-Non Military
23. High Reliability Source
24. Low Reliability Source
25. Conditional Phrases
26. Unknown Information
27. Other

III. FACTOR AND CLASS DEFINITIONS

Factor Definitions

A. Movement and Capability for Movement--relates to various aspects of movement of vehicles and men and the capability for such movement. Examples of this type of information are: the laying and support capability of pontoon structures, the actual movement of a troop unit or an aggressor vehicle, or the speed of such movement and the capability for amphibious movement.

B. Numbers, Trends, and Source--refers to the numerical aspects of units, supplies, personnel, events, etc. and trends in such aspects (e.g., number, volume increases or decreases). This factor also includes data related to the reported source of information.

C. Type of Activity--includes descriptions of the character of various types of operations, such as: reconnaissance, patrol, sabotage, and exchange of fire.

D. Training, Morale, and Attitudes--deals with the training, motivation/morale, and attitudes of the hostile forces and with information which bears on these. Examples are: the disclosure of courts martial, officer training level, and special training.

E. Communications--includes information related to field communications and the capability for such communication.

F. Armament Preparation--includes information concerning preparation of such items as: fortifications, heavy gun emplacements, mine fields, and related construction.

G. Small Arms Capability--deals with small caliber weapons, including their presence, strengths, and weaknesses.

H. Strength and Location of Units--includes information concerning strength, striking power, and location of units of all types. Examples of these types of information are: the presence of influenza in the aggressor forces, arrival of troop replacements, type and number of support units, and locational coordinates.

I. Plans and Expectations--a grouping of various types of information concerning plans, intent, and expectations. This class includes, for example, commanding officer statements, expected enemy actions, and related preparation.

J. Extra Military Information Sources--includes information from non-military and possibly unreliable sources of information. Typifying this factor are such sources as civilians and civil broadcasts.

K. Deserter Information--concerned solely with deserter related information. Examples are: deserter claims and actions.

L. Logistics--includes information relative to movement of supplies (other than armament) and equipment availability.

M. Unique Actions--includes information about those actions which are so out of the ordinary that they command special attention. An example is the entry of the Commanding General into the area.

N. Contents of Documents--concerned with information derived from written materials. An example is information derived from captured documents.

O. Civilian-Military Relationships--solely concerned with information regarding the interactions and relationships of civilians and the military. An example is the cooperation between the population and the aggressor.

Class Definitions

1. Composition of Units--the names and composition of aggressor units and temporary attachments of the various types of units (e.g., mechanized, artillery, engineering).

2. Markings and Insignia--the markings or insignia of all the various types of military units and items.

3. Position of Forces and Equipment--the specific location of all types of units, fortifications, and installations.

4. Strength of Units--the absolute and relative strength of units including number and type of equipment and number of men.

5. Replacements--the receipt, source, and deployment of personnel replacements.

6. Quantitative Indications--numerical or related quantitative statements. Examples of such statements are: 30% of, little, rest of, many, per day, a high percentage of, poorly, to the fullest, and easily.

7. Weapons and Equipment--includes information about types, characteristics, sizes, issue, supply, maintenance, and deployment of weapons and equipment.

8. Analysis--concerns detailed analyses of various types. Included in this class is information such as analyses of: reports, data, aggressor actions, etc.

9. Characterization of Missions--the nature of various missions (e.g., defensive, support, reconnaisance) and directions, requests and statements relative to current and planned action.

10. Sabotage--sabotage or the effects of sabotage efforts on the part of either friendly or aggressor forces.

11. Actual and Threatened Contact--threatened and actual contact and potential contact that did not materialize. Examples of the types of information included in this class are: exchange of small arms fire, and whether or not contact is made.

12. Strategy and Tactical Doctrines-- statements about aggressor tactical doctrine (such as, emphasis on maneuverability).

13. Preparations and Readiness--contains information concerning capability and preparedness.

14. Movement--concerned with movement of all types of units. Included in this class is such information as: direction and location of movement of convoys of various types.

15. Chronolog--concerned with all statements which refer to time of day, date, or duration. This class includes information about specific times and dates as well as general statements about time (such as, in the past, and nightly).

16. Information about Officers--includes information relevant to aggressor officers. This includes names and backgrounds of individual officers as well as other assorted information (such as disturbances in an officer's personal life).

17. Competence of Personnel and Equipment--training and experience of personnel and current capabilities and training of units and individuals.

18. Mental and Physical Condition of Troops--morale, indications of morale and health conditions of both the aggressors and civilians.

19. Civilian Military Relations--contact and relations between the military and the civilian population. Depending upon the nature of the military-civilian contact, this information may or may not also be classified under the factor Civilian-Military Relationships (factor O). For example, "population cooperates with aggressor," would be included in this class and would also be included under the Civilian-Military Relationships. However, "patrol questions residents," would be classified in both this class and in the factor Type of Activity (factor C).

20. Sources/Methods of Obtaining Information--the sources of information and the methods of obtaining information. For example, interrogation of an aggressor sabotage team, deserter based reports, sightings of vehicles and enemy activity and conversations that were overheard.

21. Informer Source-Military--deals with intelligence sources such as deserters and POWs.

22. Informer Source-Non-Military--deals with civilian intelligence sources.

23. High Reliability Source--intelligence sources such as: radar based reports, ground surveillance equipment data, and reports from U.S. units.

24. Low Reliability Source--intelligence sources such as: civilian radio broadcasts and rumor.

25. Conditional Phrases--phrases such as: since, only, when, as a result of, after, and indicates.

26. Unknown Information--refers to information that lacks some relevant facts (such as unit designations, destination, mission, or location).

27. Other--items that do not seem to fit into any other class. Examples of this are: negation, references to causes (e.g., political unrest), and unusual words (e.g., commercially and telephone).

IV. MODEL CATEGORIZATION OF PHRASES

This compendium of categorized phrases is an example of the proper use of the system. Key phrases were derived from a set of battlefield messages and classified by an expert in the use of this system. A complete list of all types of information that might possibly be stored in a given category is not included. Rather, an example of what resulted from the application of the system to one set of messages is included. This model should be used to clarify, when necessary, the type of information that may be included in a given category.

The model was developed as part of a linguistic analysis of battlefield language. The messages were broken up into two different types of meaningful phrases. Both types are seen in the pages that follow. To clarify the interpretation of these two types of phrases, consider these examples:

type 1 - in the broadcast
type 2 - promise (broadcasts, safe passage and assylum)

The first type of phrase is read as usual. The second type of phrase must have the words rearranged before being read, as follows: "broadcasts promise safe passage and assylum." The pattern can be explained using a shorthand notation substituting letters for words. A phrase written in the form "a(b,c)" is read "bac." In general, the whole phrase must be considered to understand the classification.

An understanding of the reasons for the use of two types of phrases is not necessary for the use of the system. Briefly described, the second type of phrase usually starts with a word that modifies the words within the parentheses, while the nature of the first type of phrase is more obscure.

Factor A. Movement and Capability for Movement

Class

1. Composition of Units by Type and by Designation

is composed of (convoy, 15 half ton trucks)
has (train, 25 cars)
are in (amphibious tanks, the force)
is loaded with (train, combat supplies and ammunition)
carries (train, closed boxcars)
is (apparent cargo, troop replacements)
is attached to (unknown mech unit, 1F div)
is attached to (an element of the eng bn, 1F mech div)
of the 1F div
within the 34th. and 35th. mech div

2. Markings and Insignia

wear (aggressor enlisted men, 35 mech div insignia)
have (14 trucks, 16 CAA bumper markings)
are (bumper markings, 210)
are marked (10 heavy tanks, 3-10)
have (trucks, 34 mech div markings)
bear (patrols, 1F mech insignia)
are designated (bridges, Tango 2)
have (trucks, 93 armored regt markings)

3. Position of Forces and Equipment

is in (1F mech div, area)
are throughout (personnel and vehicles, the area)
is in (1F mech div, this location)
are at (pontoon sections, vic UR 048900)

4. Strength of Units

is (the personnel strength of the 35 mech div, 70-80% TO&E)
is (strength of 547 mech regt, 78% of TO&E)

7. Weapons and Equipment

are heavy (pon sections)
support (sections, 60 tons)
is (unit, amphibious)
deploys (210 rgt pontoon sections)
are laid (pontoons)
are equipped with (some tank units, fording equipment)

Factor A. Movement and Capability for Movement

(continued)

Class

8. Analysis

analyzes (corps, aggressor ammunition movement)

9. Characterization of Missions

is (mission, to support major river crossing)
1F mech div requests that

10. Sabotage

is sabotaged (aggressor rail system)

11. Actual and Threatened Contact

did not cross (aggressor aircraft, border)
has mission of attacking (rgt 20 mech div)

12. Strategy and Tactical Doctrine

stresses (aggressor tactical doctrine, speed and shock action)

13. Preparations and Readiness

prepare (aggressor forces of 35 mech div, bridges)

14. Movement

move (trucks, NE)
moves (train, south)
move (4 large tk trucks, north)
move (troops, SE)
move to (regt, fwd assembly area)
moves (aggressor unit, into territory)
occurs (crossing into friendly territory)
occurs (forward displacements)
comes in (traffic)
heads (convoy, NE VIC US 155095)
leaves (train, AUE (US 3706))
heads (train, SW VIC UR 115955)
pass through (aggressor units, town of Pfaffengrun)
pass through (two groups of wheeled vehicles)
pass (one large group tks or SP guns)
is indicated (a large movement of troops)

Factor A. Movement and Capability for Movement

(continued)

Class

16. Information about Officers

visits (G/A Mikhaylov, maneuver area of 34 mech div)

17. Competence of Personnel and Equipment

is ranked as (1F mech div, excellent)
is trained (1F mech div)

20. Sources/Methods of Obtaining Information

have been sighted (mech units)
are sighted (both support and attack type aircraft)
is from (deserter, 2/25 mech regt 35 mech div)
comes (deserter, to our lines)

26. Unknown Information

is unknown (destination)

27. Other

by air

Factor B. Numbers, Trends and Source

Class

1. Composition of Units by Type and by Designation

   is composed of (convoy, 15 half ton trucks)
   has (train, 25 cars)
   consists of (patrol, 20 men)
   are used (400 additional civilian workers)
   receive (aggressor units, 300 men)

2. Markings and Insignia

   have (14 trucks, 16 CAA bumper markings)
   are marked (10 heavy tanks, 3-10)

3. Position of Forces and Equipment

   are in (major aggressor units, the following areas)
   are in (major aggressor units, Schoneck)
   are in (2 aggressor bns, Schoneck)
   is in (large storage area, VIC UR 032981)

5. Replacements

   receives (16 CAA, 160 men)

6. Quantitative Indications

   percent of
   a large number of
   fewer than
   many of the
   per day
   little
   large volume of

7. Weapons and Equipment

   are new (weapons)
   is defective (equipment)

11. Actual and Threatened Contact

   is made with (significant contact, friendly units)
   is minor (contact)

Factor B. Numbers, Trends and Source

(continued)

Class

14. Movement

pass (two groups of wheeled vehicles)
pass (one large group tks or SP guns)
is indicated (a large movement of troops)

15. Chronology

nightly
continually
continuing for ½ hour
for the past 96 days
beginning 130240 sep

18. Mental and Physical Condition of Troops

arrested (MPs, 12 aggressor soldiers)

20. Sources/Methods of Obtaining Information

have been sighted (mech units)
are sighted (both support and attack type aircraft)
is from (deserter, 2/25 mech regt 35 mech div)
is from (deserter, 210 eng regt)

Factor C. Type of Activity

Class

1. Composition of Units by Type and by Designation

   is composed of (convoy, 15 half ton trucks)
   consists of (patrol, 20 men)

2. Markings and Insignia

   bear (patrols, 1F mech insignia)

3. Position of Forces and Equipment

   is in (jamming unit, 1F mech div area)

4. Strength of Units

   is implemented (rationing)

7. Weapons and Equipment

   is fired (weapon)
   had (the 1F arty div, misfires)
   use (aggressor troops, weapons)
   maintain (aggressor troops, weapons)
   deploys (210 regt, pontoon sections)
   are equipped with (some tank units, fording equipment)

8. Analysis

   analyzes (corps, aggressor ammunition movement)
   has been analyzed (recent history of the 16 CAA)

9. Characterization of Missions

   is (mission, to support major river crossing)
   is (mission, infared photography)
   is being carried out

10. Sabotage

    is sabotage (the aggressor rail system)
    performs sabotage (dissidents within the official party)
    destroys (aggressor sabotage, POL supplies)
    infiltrates (aggressor sabotage team, US lines)
    creates (aggressor sabotage team, confusion)
    jams (aggressor, transmission)

Factor C. Type of Activity

(continued)

Class

11. Actual and Threatened Contact

has mission of attacking (rgt 20 mech div)
is made with (significant contact, friendly unit)
is minor (contact)
is made (contact)
is made (no contact)
have contact with (US units, aggressor patrols)
withdrew (enemy patrols)
disengaged (aggressor)
occurs (combat)
occurs (a short (3 min) exchange of small arms fire)
provokes (US, aggressor)

13. Preparations and Readiness

did emplacement of (work force, aggressor mine fields)
prepare (civilian workers, fortifications)
are being prepared (heavy gun emplacements)
are prepared (positions)
occurs (preparation)
prepare (aggressor forces of 35 mech div, bridges)

14. Movement

are major (logistical movements)

15. Chronology

during contact
for the duration of exercise

16. Information about Officers

reviews (CG, progress of maneuvers)

18. Mental and Physical Condition of Troops

arrested (MPs, 12 aggressor soldiers)

19. Civilian-Military Relations

relocates (unit, some civilian residents of Warschitz and Leubetha)
questions (patrol, residents)

Factor C. Type of Activity

(continued)

Class

20. Sources/Methods of Obtaining Information

have been sighted (mech units)
intercepts (unit, radio traffic)
are observed (firings)
is sighted (aggressor patrol)
are in (intelligence sources, aggressor area)
was interrogated (aggressor-sabotage team)
was captured (deserter, VIC TR 9496)
was captured (aggressor sabotage team)
picks up (patrol, deserter)

27. Other

under blackout conditions

Factor D. Training, Morale and Attitudes

Class

5. Replacements

are trained (replacements)
are from (replacements, satellite nations)

7. Weapons and Equipment

are superior (US forces)

16. Information about Officers

reviews (CG, progress of maneuvers)
graduated from (G/D Borodin, Alta Komando College)
are from (eng officers, 16 CAA staff)
is depressed (Turgenev)
is well versed in (G/D Borodin, urban conflict)

17. Competence of Personnel and Equipment

are trained in
is trained (1F mech div)
is rated as (1F mech div, excellent)
have (replacements, technical qualifications)
are competentin (most officers and enlisted men, individual combat)
are competent in (most officers and enlisted men, weapons operation)
are competent in (most officers and enlisted men, CBR warfare)
are able (generals and marshalls)
exists (capability for large scale operational support)
proves (the exercise, some major new concepts)
have (officers, high level staff experience)
have graduated from (senior NCOs, aggressor, NCO academy)
are trained in (units, urban combat)
are apparent (training and discipline)

Factor D. Training, Morale and Attitudes

(continued)

Class

18. Mental and Physical Condition of Troops

arrested (MPs, 12 aggressor soldiers)
broke out (a fight)
may cause (integration, problems)
are confused (people)
is declared (martial law)
is good (morale)
have been disclosed (several courts martial)
has increased (abuse pilferage)

19. Civilian-Military Relations

cooperates with (population, aggressor)
relocates (unit, some civilian residents of Warschitz and Leubetha)
cooperate with (civilians, military)
cooperate with (political leaders, aggressor military)
is apprehensive of (population, nuclear weapons)

27. Other

due to the recent political unrest
because of political beliefs

Factor E. Communications

Class

10. Sabotage

    jams (aggressor, transmissions)

13. Preparations and Readiness

    alerts (signal, 1F mech div)

18. Mental and Physical Condition of Troops

    promise (broadcasts, safe passage and asylum)
    is poor (aggressor communication discipline)

20. Sources/Methods of Obtaining Information

    were overheard (conversations)
    are from (reports, aggressor homeland)
    intercepts (unit, radio traffic)
    is intercepted (signal)
    are monitored (radio broadcasts)

26. Unknown Information

    is unintelligible (message)

27. Other

    by telephone

Factor F. Armament Preparation

Class

7. Weapons and Equipment

are in (aggressor-mine fields, patterns)
are bunkered (positions)

13. Preparations and Readiness

did emplacement of (work force, aggressor mine field)
prepare (civilian workers, fortifications)
are being prepared (heavy gun emplacements)
are prepared (positions)
prepare (aggressor forces of 35 mech div, bridges)

16. Information about Officers

supervise (eng officers, emplacement)

Factor G. Small Arms Capability

Class

7. Weapons and Equipment

has been increased (range of rifle)
is important (new rifle)
is known about (little, new rifle)
are similar to (weight and size, US M-16)

Factor H. Strength and Location of Units

Class

1. Composition of Units by Type and Designation

   is composed of (1F mech div, 211 MDM TK RGT)
   is composed of (1F mech div, 208 mech rgt)
   is composed of (1F mech div, 211 mech regt)
   is composed of (1F mech div, 214 mech rgt)
   is in (unit, 211 rgt)
   is (apparent cargo, troop replacements)
   consists of (patrol, 20 men)
   are used (400 additional civilian workers)
   receive (aggressor units, 300 men)

2. Markings and Insignia

   have (14 trucks, 16 CAA bumper markings)
   are marked (10 heavy tanks, 3-10)

3. Position of Forces and Equipment

   is in (1F mech div, area)
   is in (54 mech div 19 CAA, vic Auerbach)
   is in (deserter's unit, UR 037847)
   is in (unit, vic Neustadt)
   is in (div, position)
   are in (major aggressor units, the following areas)
   are (aggressor units, further north)
   are in (major aggressor units, Schoneck)
   is in (jamming unit, 1F mech div area)
   are in (two agressor bns, Schoneck)
   are throughout (personnel and vehicles, the area)
   is in (1F mech div, this location)
   is north of (aggressor, three corps area)
   is at (engineering equipment, this vicinity)
   are at (pontoon sections, vicinity UR 048900)
   is in (large storage area, vic UR 032981)

4. Strength of Units

   is (the personnel strength of the 35 mech div, 70-80% TO&E)
   is unit strength, 83% of TO&E)
   is (strength of 547 mech regt, 78% of TO&E)
   requires (strength, attention)
   is implemented (rationing)
   are strong (major or subordinate units)
   are reduced (unit strengths)

Factor H. Strength and Location of Unit

(continued)

Class

5. Replacements

receives (16 CAA, 160 men)
receive (units, replacements)
is requisitioning (16 CAA, replacements)
receive (aggressor units, replacements)
have been replaced (personnel)
will be integrated into (replacements, 16 CAA units)

6. Quantitative Indications

a highly resistant strain of

7. Weapons and Equipment

are equipped with (some tank units, fording equipment)

8. Analysis

is analyzed (aggressor strength)

14. Movement

move (4 large tk trucks, north)
pass (one large group of tks or SP guns)

16. Information about Officers

are in (generals and marshalls, the aggressor army)
are in (officers, aggressor ground forces)
is assigned to (aggressor major, 16 CAA general staff)
is (I.R. Presnosff, BN CDR)
is (Col V Weir, Rgt CDR)
was replaced (deserter's unit's CDR)
is CG of (G/A Mikhaylor, D., 19 CAA)

18. Mental and Physical Condition of Troops

have (hospitals, high level of activity)
infects (influenza, several units)

Factor H. Strength and Location of Unit

(continued)

Class

20. Source/Method of Obtaining Information

are sighted (both support and attack type aircraft)

26. Unknown Information

does not know (deserter, location of unit)
are not determined (unit designations and strength)
are not reported (hospital locations)

Factor I. Plans and Expectations

Class

5. Replacements

   is requisitioning (16 CAA, replacements)

9. Characterizations of Missions

   have (military uses, priority)
   is (mission, to support major river crossing)
   is on defensive (16 CAA)
   will be carried out (assigned missions)
   has mission of
   intend to
   purpose is to
   for purpose of
   with purpose
   their mission is
   the 16 CAA authorizes
   1F *mech DIV requests that*
   signal instructs
   enemy permits

11. Actual and Threatened Contact

   has mission of attacking (rgt 20 mech div)
   exists (a threat to the border)

12. Strategy and Tactical Doctrine

   stresses (aggressor tactical doctrine, speed and shock action)
   is ramified (tactical doctrine)
   emphasizes (aggressor tactical doctrine, fire and maneuver)
   is emphasized (combat power of the aggressor)

13. Preparations and Readiness

   are being prepared (heavy gun emplacements)
   are prepared (positions)
   occurs (preparation)
   alerts (signal, 1F mech div)
   is prepared (div)

Factor I. Plans and Expectations

(continued)

Class

14. Movement

move (aggressor unit, into positions)

15. Chronology

soon
immediately
beginning 130240 Sep
for early Sep

16. Information about Officers

is on (CG G.S. Kuvov, offensive)
emphasizes (G/O Borodin, the use of infantry)

17. Competence of Personnel and Equipment

are competent in (most officers and enlisted men, CBR warfare)
exists (capability for large scale operational support)

18. Mental and Physical Condition of Troops

may cause (integration, problems)
promise (broadcasts, safe passage and asylum)

Factor J. Extra-Military Information Sources

Class

18. Mental and Physical Condition of Troops

promise (broadcasts, safe passage and asylum)

20. Sources/Methods of Obtaining Information

were overheard (conversations)
were interrogated (civilian line crossers)
are from (reliable civilian informers, East Planen)

22. Information Source-Non Military

civilians report
civilian reports from 34 div area indicate
friendly civilians indicate
civilians indicate
friendly civilians overhear
friendly civilians verified
civilian reports indicate
reliable civilian informers report
civilians sight
rumers continue within the civilian community

27. Other

in the broadcast

Factor K. Deserter Information

Class

3. Position of Forces and Equipment

   is in (deserter's unit, UR 037847)

16. Information about Officers

   was replaced (deserter's unit's CDR)

20. Sources/Methods of Obtaining Information

   is from (deserter, 2125 mech regt 35 mech div)
   has instructed (commander, deserter)
   is from (deserter, 210 eng regt)
   bases claim on (deserter, observation)
   comes (deserter, to our lines)
   picks up (patrol, deserter)

21. Information Source--Military

   deserter states that
   deserter stated
   deserter claims that
   deserter reports
   deserter indicated

26. Unknown Information

   does not know (deserter, location of unit)

Factor L. Logistics

Class		
1.	Composition of Units	
		is loaded with (train, combat supplies and ammunition) is (apparent cargo, supplies)
7.	Weapons and Equipment	
		deploys (210 regt., pontoon sections)
10.	Sabotage	
		is sabotaged (the aggressor rail system)
14.	Movement	
		are major (logistical movement)
26.	Unknown Information	
		is unknown (destination)

Factor M. Unique Actions

Class

7. Weapons and Equipment

   employ (units, CBR)

9. Characterization of Missions

   CG 1F orders

11. Actual and Threatened Contact

    creates (US, international incident)

16. Information about Officers

    visits (G/A Mikhaylov, maneuver area of 34 mech div)
    reviews (CG, progress of maneuvers)
    is on (CG G.S. Kurov, offensive)
    is CG of (G/A Mikhaylov, area of 34 mech div)

17. Competence of Personnel and Equipment

    are competent in (most officers and enlisted men, CBR warfare)
    proves (the exercise, some major new concepts)

20. Sources/Methods of Obtaining Information

    has instructed (commander, deserter)

Factor N. Contents of Documents

Class

8. Analysis

are analyzed (data and reports)
are available to (data and reports, G-2 3 corps)

9. Characterization of Missions

aggressor document stresses the requirement for

20. Sources/Methods of Obtaining Information

is intercepted (document)
obtain (intelligence sources, document)
is captured (document)
obtains (civilian, document)

Factor O. Civilian and Military Relations

Class

13. Preparations and Readiness

prepare (civilian workers, fortifications)

19. Civilian-Military Relations

cooperates with (population, aggressor)
relocates (unit, some civilian residents of Warschitz and Leubetha)
cooperates with (civilians, military)
cooperates with (political leaders, aggressor military)
questions (patrol, residents)
are within (civilians, 1F mech area)
is apprehensive of (population, nuclear weapons)

20. Sources/Methods of Obtaining Information

were interrogated (civilian line crossers)
obtains (civilian, document)
are from (reliable civilian informers, East Plauen)

22. Informer Source--Non-Military

rumors continue within the civilian community

V. MODEL: ISONYMS BY CATEGORY

The user of this system should be alerted to the imprecise use of words in the battlefield language. Listed here are words which are used synonomously within categories. This list was derived from the messages used in the model in Section IV and is not exhaustive.

Category	Isonyms	
A1	loaded	composed of
	carries	attached
	cargo	
A2	wear	
	bear	
	have	
	marked	
	designated	
	insignia	
A7	deploy	
	lay	
A14	moves	
	heads	
B1	composed of	
	consists of	
C1	composed of	
	consists of	
C20	sighted	
	observed	
	captured	
	picked up	
L1	loaded with	
	cargo	
N20	intercepted	
	obtained	
	captured	
D17	trained in	
	competent in	
	are able	
D27	due to	
	because	
E20	intercepted	
	monitored	

Category	Isonyms
H26	does not know are not determined
I9	mission intends purpose
I12	stress emphasize
J22	reports indicates
K21	stated claims reports indicates

VI. MODEL: WORD DICTIONARY

To further clarify the use of this system, a dictionary is presented here consisting of all the words used in the model in Section V. If you are unsure of the categorization of a given message, you may choose a key word from the message and look it up in this dictionary. This will tell you in which categories that word is to be found in the model. This will not work every time since the classification is based on the meaning of the whole message, not any one word. However, this method may point to categories you had not considered previously.

APPENDIX C

Derivation of Extension of Fleiss' Technique for More than One Assignment

PRECEDING PAGE BLANK-NOT FILMED

A Measure of Agreement Among Several Raters with Each Making Several Assignments

Part of the method given by Fleiss (1971) for measuring agreement among several judges, each making only one assignment, is extended to each making more than one assignment. The new method applies a correction to each individual message to account for the possibility of guessing.

Let k = number of categories available for assignment to a message
n = number of judges
n_j = number of judges who assigned category j
a = number of categories assigned to a message by each judge

A raw measure of agreement of assignments to a particular message is given by:

$$P = \frac{\sum_{j=1}^{k} n_j(n_j-1)}{an(n-1)}$$

For a = 1, this agrees with Fleiss. The agreement, corrected for the possibility of guessing, is given by:

$$A = \frac{P - P_r}{1 - P_r}$$

where: P_r is the P obtained by assuming the most random results possible for n judges, i.e., the assignments (categorizations) are most evenly divided among the k categories.

PRECEDING PAGE BLANK-NOT FILMED

Example I

As an example, let k = 60 categories, n = 15 judges, and a = 6 assignments per judge. There are thus 15 x 6 = 90 assignments distributed among the 60 categories or an average of 1.5 assignments per category. Therefore, for the most even distributions some categories will have two assignments and some one. Let x be the number of categories having two assignments. Then, there will be 60 - x categories with one assignment. The total number of assignments is $2x + 1(60 - x) = 90$, which gives: $x = 30$. Therefore, the most random assignment would give 30, which have agreements of two judges and 30 single assignments.

Since for $n_i = 2$, $n_i(n_i - 1) = 2$, and for $n_i = 1$, $n_i(n_i - 1) = 0$, P_r has 30 twos in the numerator.

$$P_r = \frac{30(2)}{6(15)(14)} = \frac{1}{21}$$

which is the result obtained from the most likely random assignment. The corrected measurement of agreement would then be:

$$A = \frac{P - \frac{1}{21}}{1 - \frac{1}{21}} \approx \frac{21P - 1}{20} = 1.05\,P - .05$$

for the parameters assumed in this example. If such a case actually occurred, A would be zero. For P = 1, A would be 1. Suppose that for a given message, there was actually one category on which all 15 judges agreed, two categories on which 12 agreed, two on which 5 agreed, and 41 categories chosen by only one judge. This gives:

$$P = \frac{15(14) + 2[12(11)] + 2(5(4)] + 41(0)}{6(15)(14)} = \frac{494}{1260} = .392$$

$$A = 1.05(.392) - .05 = .362$$

Suppose for another message, there were two categories on which 14 judges agreed, three on which 11 agreed, seven on which 3 agreed, and four on which 2 agreed.

$$P = \frac{2[14(13)] + 3[11(10)] + 7[3(2)] + 4[2(1)]}{1260} = \frac{744}{1260} = .591$$

$$A = 1.05(.591) - .05 = .57$$

Example II

Let k = 40 categories, n = 5 judges, and a = 3 assignments per judge. This gives a total of 15 assignments distributed among 40 categories. The most probable random assignment would then have no agreement among the judges, i. e., 15 of the 40 categories would be chosen once.

$$P_r = \frac{15(0)}{3(5)(4)} = \frac{0}{60} = 0 \text{ and}$$

$$A = \frac{P - 0}{1 - 0} = P$$

Suppose there was one category on which all 5 judges agreed, two on which 3 judges agreed, and the other assignments were scattered.

$$A = P = \frac{5(4) + 2[3(2)] + 4(0)}{60} = \frac{32}{60} = .533$$

APPENDIX D

User Syntax Diagrams for Computer Implementation

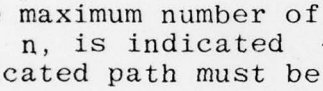
PRECEDING PAGE BLANK-NOT FILMED

Appendix

The syntax diagram shown on the following pages was chosen as a means to display the syntax which the user (intelligence analyst) of the system would employ. This type of diagram affords a concise composition of syntax involving details, alternatives and iterations which is rigorous without being cumbersome.

The basic rule in reading these diagrams is that any path traced along the forward direction of the arrows will produce a syntactically correct component of the user language. Those items enclosed by < > indicate elements which are explained by the words bracketed, i.e., not further diagramatically defined.

Words which are entirely underlined must be completely specified by user keyboard entry. Words which are partially underlined can be abbreviated by the underlined part or any part containing at least the underlined part. Accordingly, INPUT can be represented by IN, INP, or INPUT but not by I or INPUTS.

The maximum number of iterations through a portion of the diagram, n, is indicated 

. An asterisk indicates that the indicated path must be taken at least once.

Messages are supplied to the system for storage using the first syntax diagram shown, i.e., INPUT. For each message input, at least the class, and factor information and the message itself must be provided. For a further syntactical definition of the class and factor (as well as the other identified) see those respective diagrams shown later.

In making an inquiry, the system user would use the INQUIRY syntax. The syntax was designed to allow a scenario such as that which follows. The analyst prepares an initial set of selection criteria describing the classes of messages he would like to examine, search or count. To do this, he enters only INQ ALL WITH followed by any one or more items in the selection description (next to last diagram) followed by one of the inquiry types. Consider the example:

```
INQ ALL WITH CL A,FAC 2, CL F ,FAC 13, KEY BRIDGE
FR 5-17-79 TO NOW, COUNT
```

This represents a request to count all messages of categories A2 or F13 containing the word BRIDGE, entered into the system after May 17, 1979.

PRECEDING PAGE BLANK-NOT FILMED

The analyst is presented with the count.

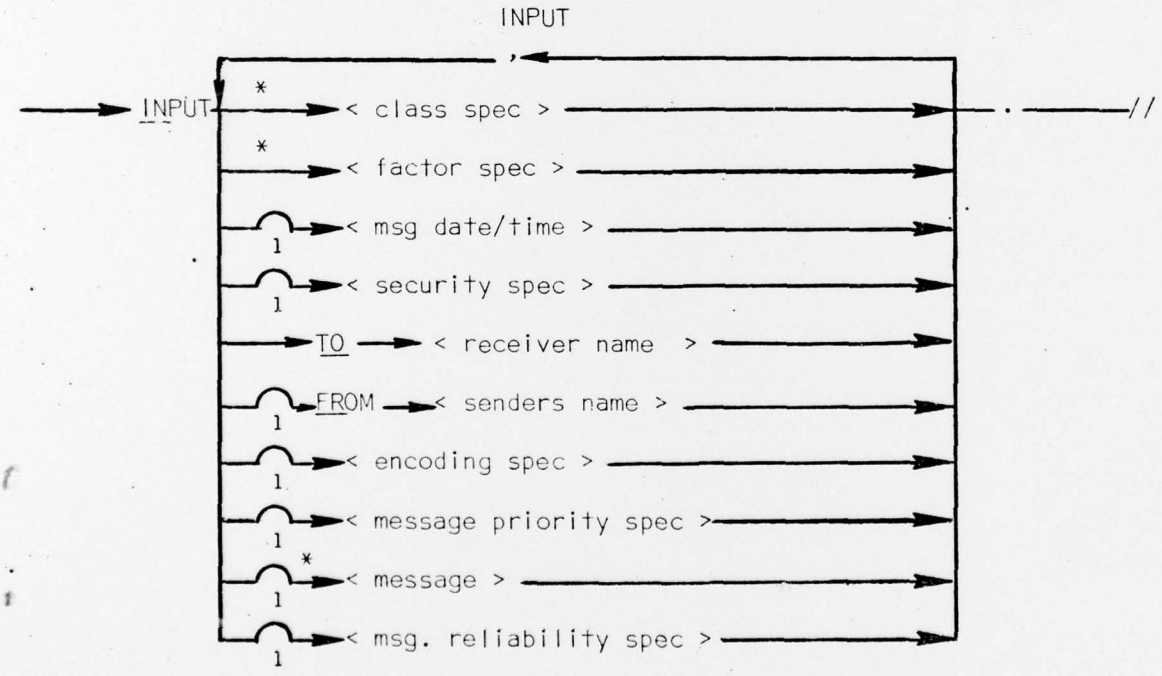
If the count is large, the analyst may enter another inquiry without the word ALL by adding information to the selection criteria. This action will reduce the number of messages to be visually inspected. For example:

INQ FROM "SENDERS NAME", COUNT

will consider only those messages which met the prior selections criteria and also have the property that they were sent from some specific sender (individual or organization by name).

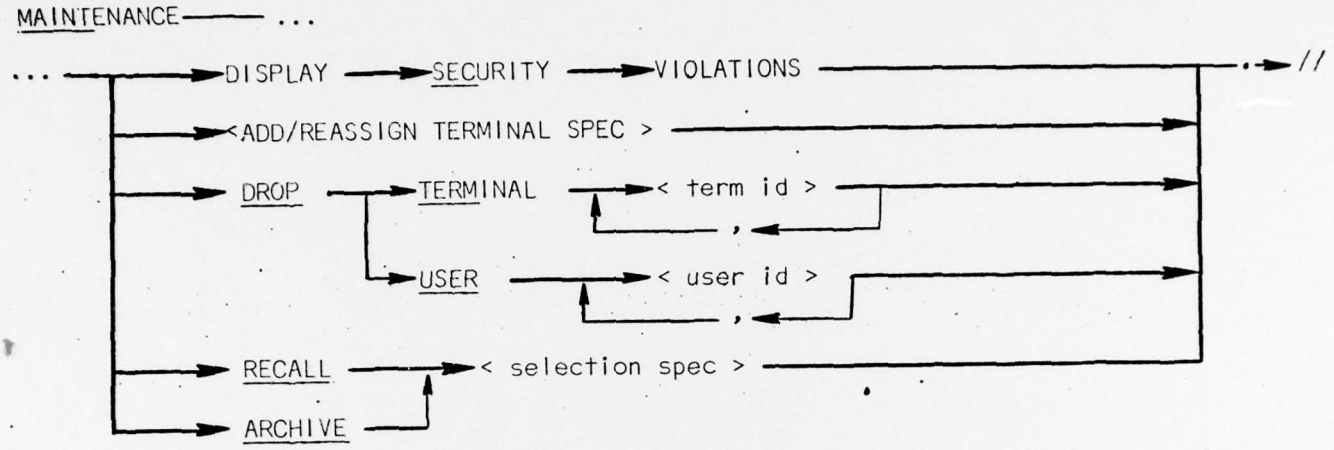
If the resultant count is sufficiently low so that the user would like to see a display of all of them, he then enters:

INQ DISPLAY
or INQ DISPLAY F (first page) followed by
INQ DISPLAY N (next page)



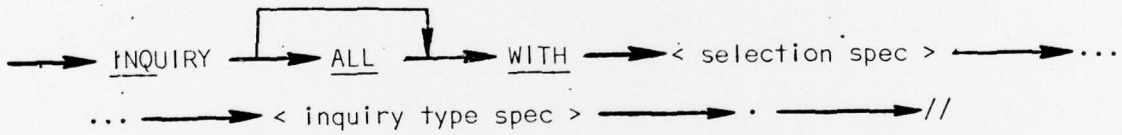
Function: provides message information to system
Default: none

MAINTENANCE



Function: controls system use
Default: no users

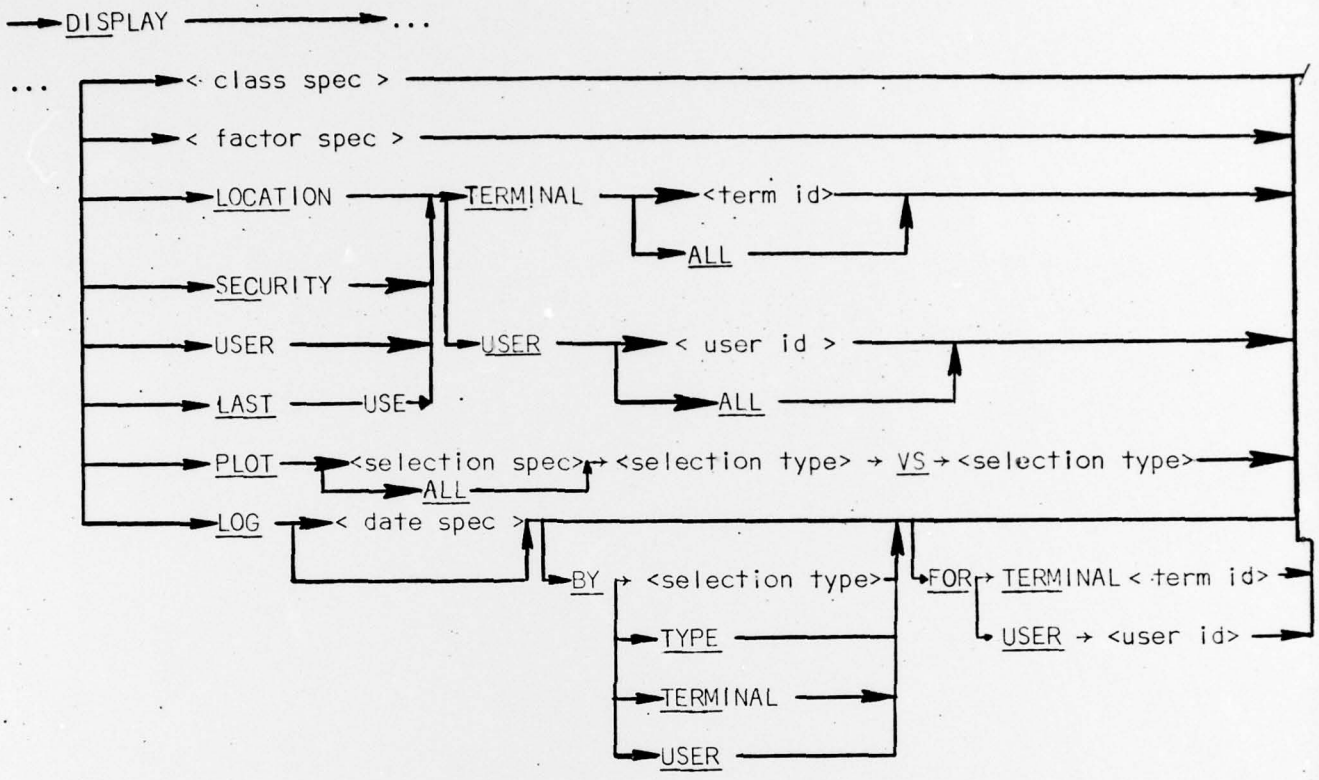
INQUIRY



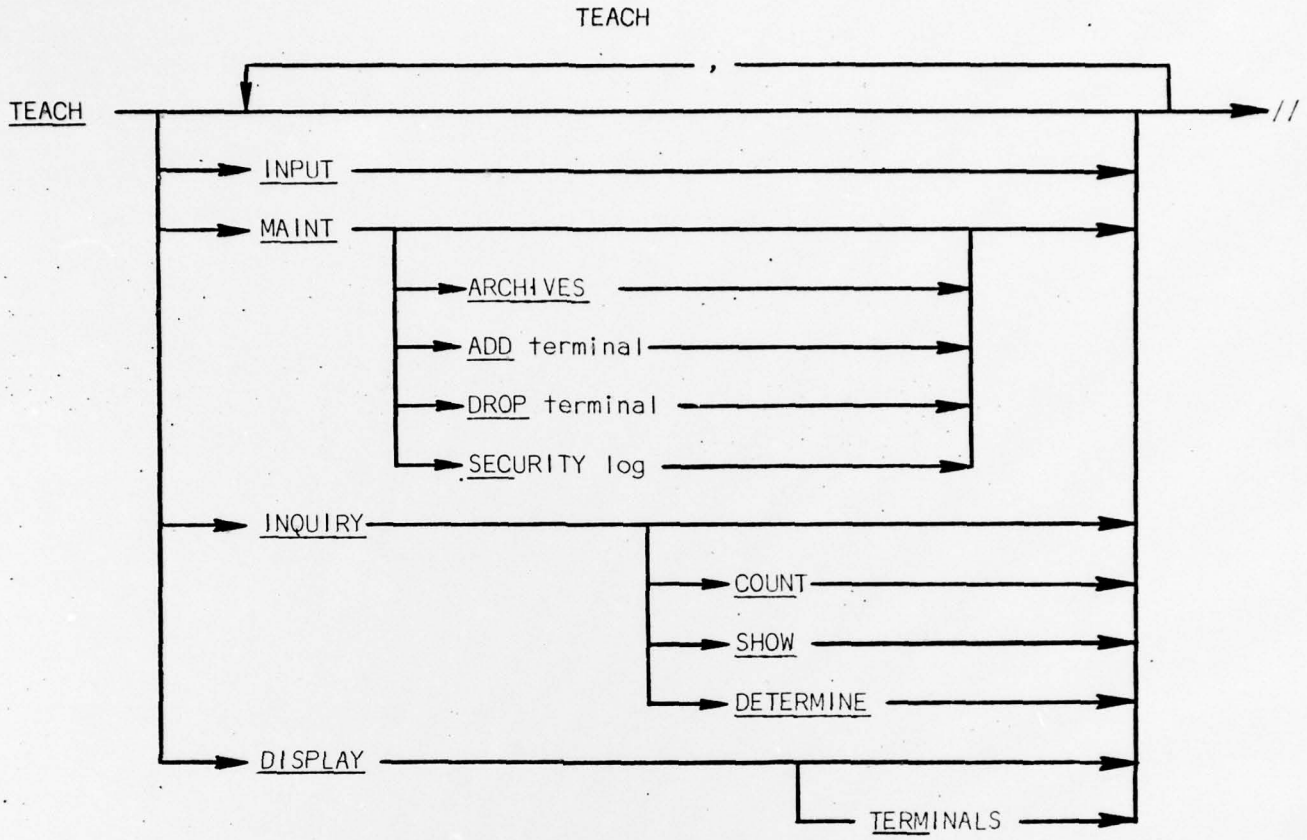
Function: asks for information about a set of messages. If the word ALL is used, then apply the < selection spec > to all messages from the set. Otherwise, apply the < selection spec > to the last formed set of messages. The information desired is specified by the <inquiry type spec >.

Default: normal configuration.

DISPLAY

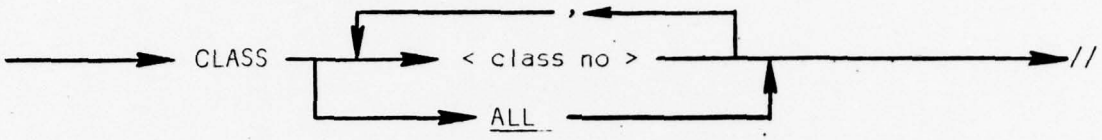


Function: present inquiry results to user
Default: normal configuration



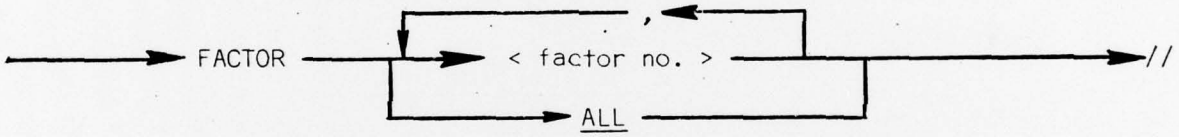
Function: user training
Default: normal configuration

< CLASS SPEC > ::=



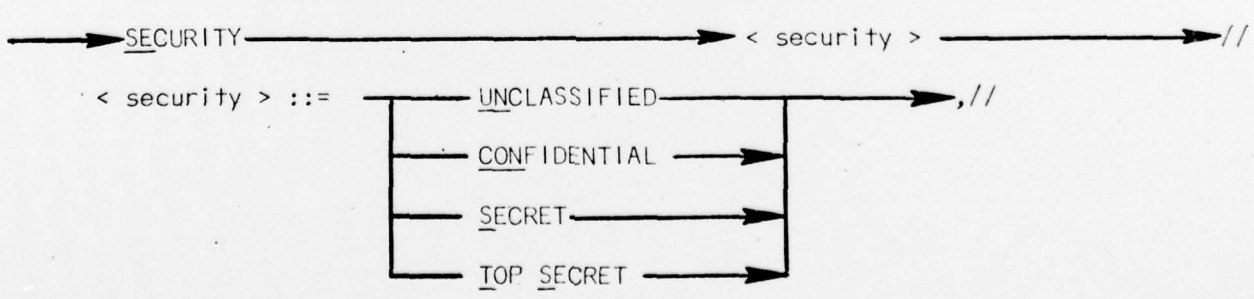
Function: specifies one or more of the 27 classes into which battlefield messages may be assigned
Default: all classes

< FACTOR SPEC > ::=



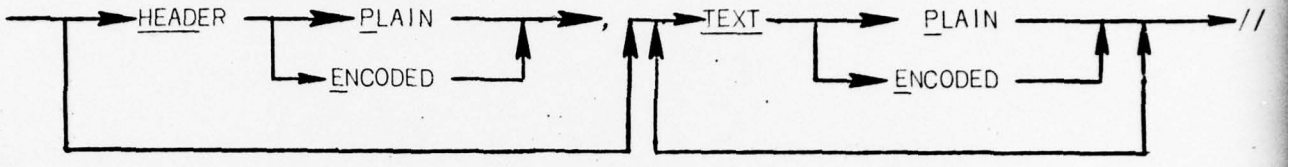
Function: specifies one or more of the 15 factors into which battlefield messages may be assigned
Default: all factors

< SECURITY SPEC > ::=



Function: defines either (1) security level of a message, or (2) highest authorized security level for a terminal
Default: unclassified security

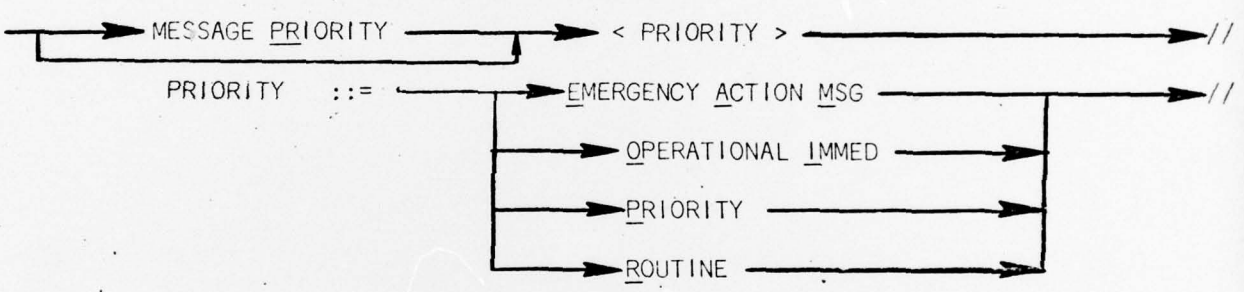
ENCODING SPEC::=



Function: specifies whether or not header and/or text is input and stored in plain or encoded form.

Default: header plain, text plain

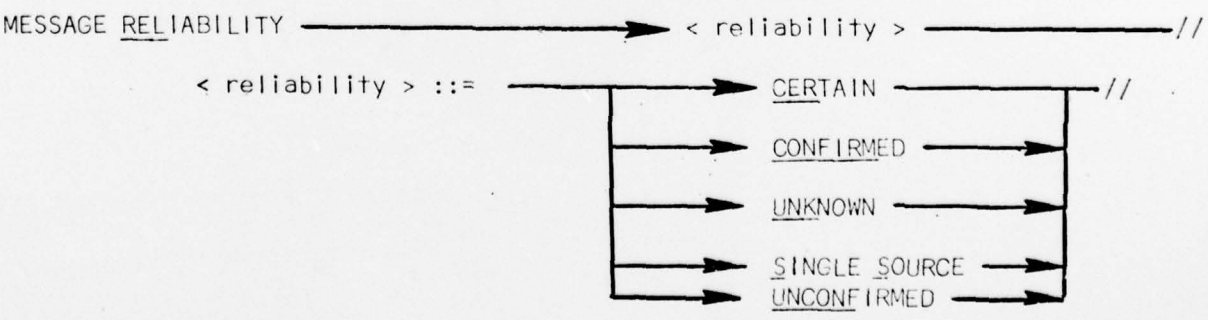
MESSAGE PRIORITY SPEC ::=



Funtion: specifies the priority/criticality of a message

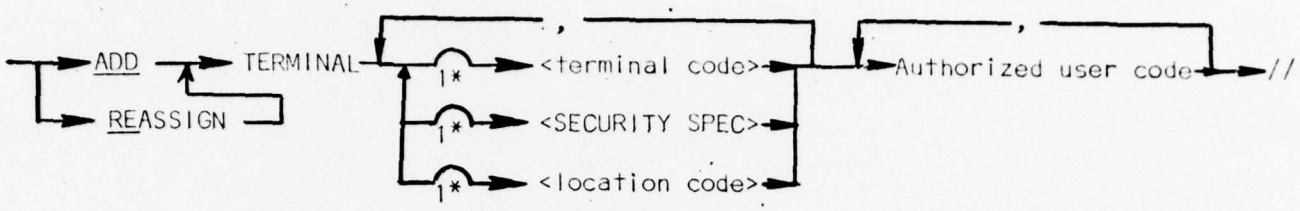
Default: message priority= routine

MESSAGE RELIABILITY SPEC ::=



Function: specifies the reliability/certainty of a message
Default: message reliability= unknown

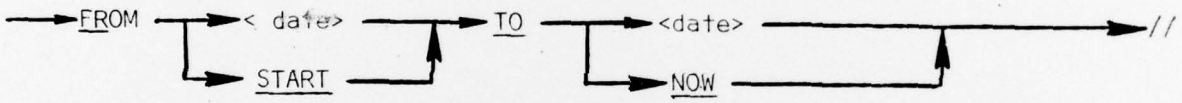
ADD/REASSIGN TERMINAL SPEC ::=



Function: to add a terminal to the network or to reassign a terminal by redefining its location, security category and/or applicable user codes

Default: none

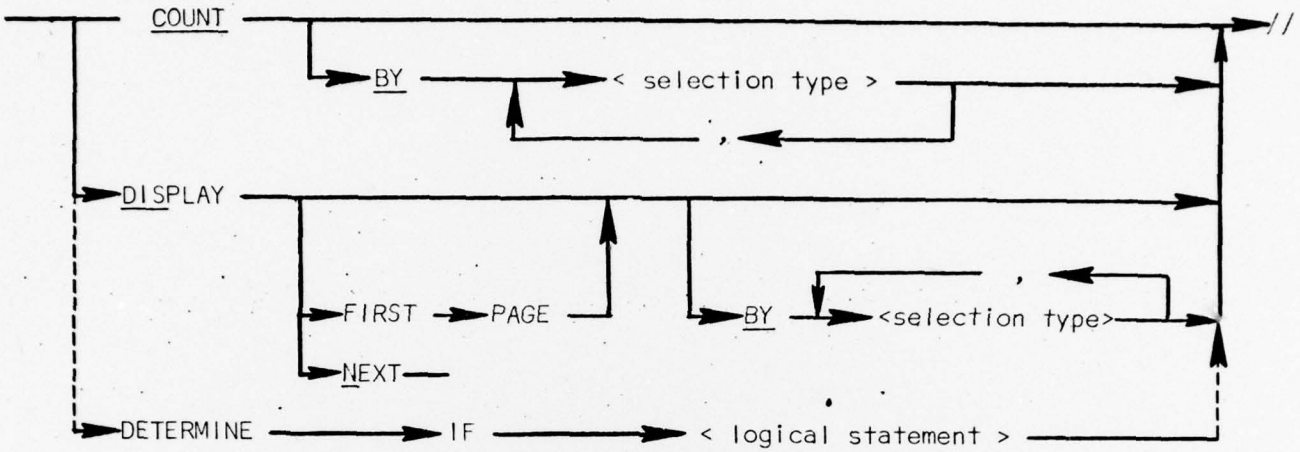
DATE SPEC ::=



Function: to define a time period over which messages of interest are to be processed.

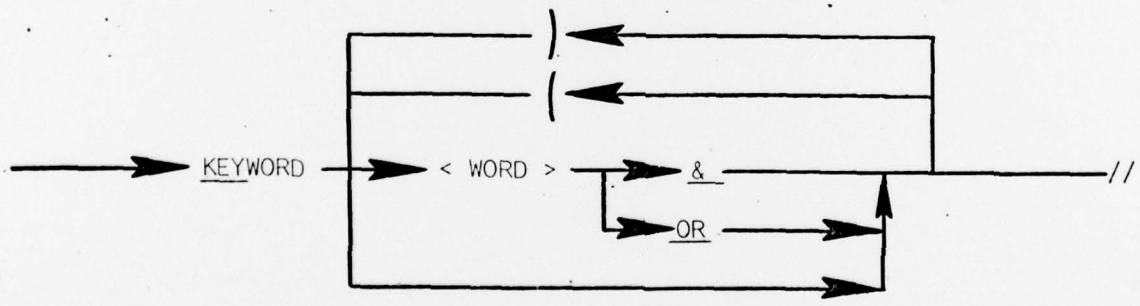
Default: from start of system to today

< INQUIRY TYPE SPEC > ::=



< logical statement > is any question that can be answered "yes" or "no" by analyzing the message's text.

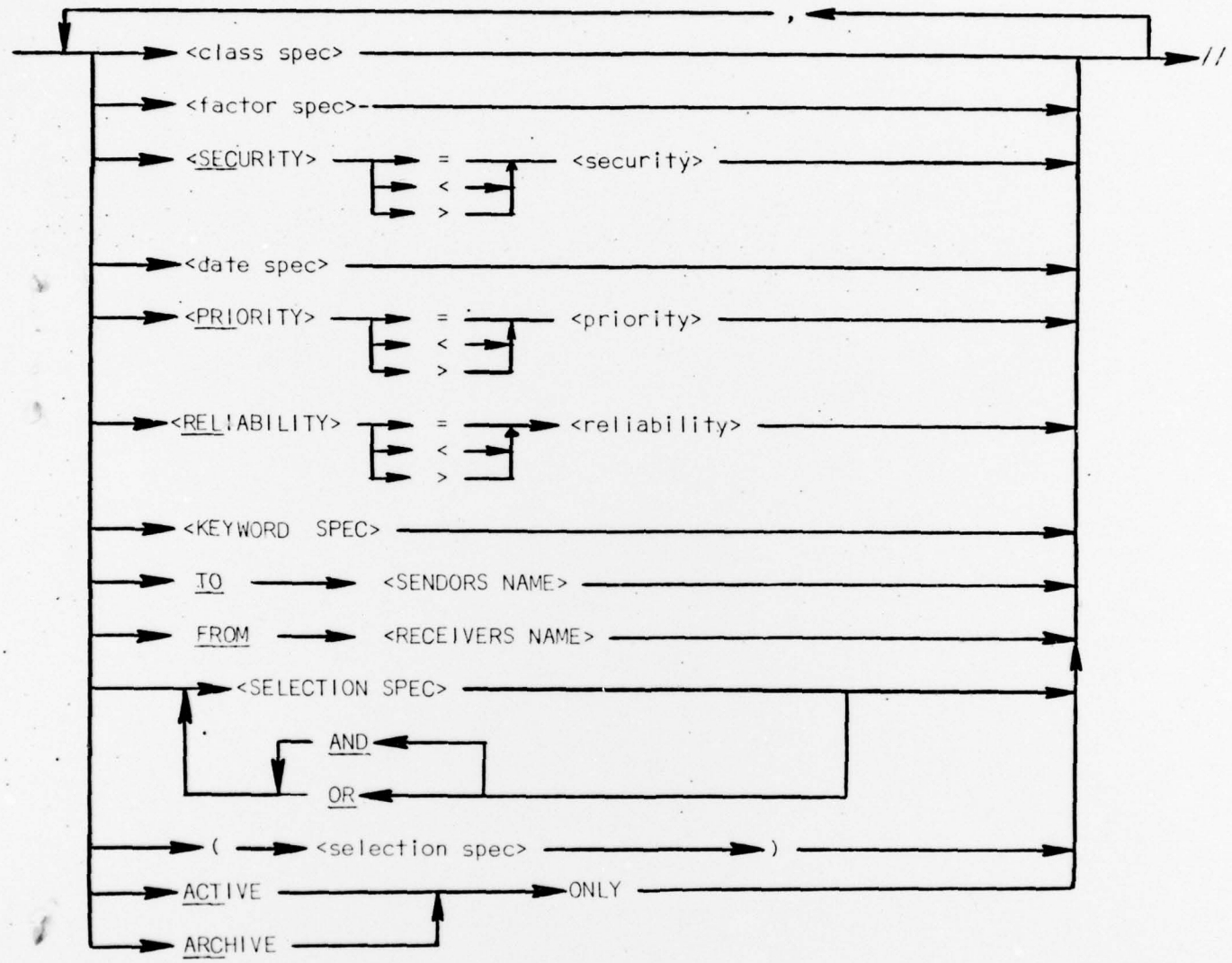
KEYWORD SPEC ::=



Function: provides the capability for user to specify one or more special words when he would like to count or retrieve stored messages containing this (these) words

Default: no key words

< SELECTION SPEC > ::=



Note:
Each <selection spec> defines a set of messages, each message having all of the specified conditions. Sets of messages can be combined by OR (messages in either or both sets) or AND (messages in both sets). Thus

PR= ROUTINE AND (SENDER X OR SENDER Y)

would be the set of messages of priority routine sent to X or Y

<SELECTION TYPE> ::=